

A/box

seg

UTILITY
PATENT APPLICATION
TRANSMITTAL

(Only for new nonprovisional applications under 37 CFR 1.53(b))

Attorney Docket No. MIT-103

First Named Inventor Kawasaki

Title Genes Integrating Signal Transduction Pathways

APPLICATION ELEMENTS

ADDRESS TO: Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 202311. ☐ Fee Transmittal Form2. ☒ Specification and Drawings [Total Pages 132]

- Specification - (108 pages)
- Claims - (20 pages)
- Abstract - (1 page)
- Sheets of Drawings - (3 sheets)
- ☐ Formal
- ☒ Informal

3. ☐ Oath or Declaration [Total Pages]

- a. ☐ Newly executed (original)
- b. ☐ Copy from a prior application (37 CFR 1.63(d))

(for continuation/divisional with Box 17 completed)
[Note Box 4 below]4. ☐ Incorporation by Reference (usable if Box 3b is checked)
The entire Disclosure of the prior application, from which a copy of the oath or declaration is supplied under Box 3b, is considered as being part of the disclosure of the accompanying application and is hereby incorporated by reference therein.5. ☐ Microfiche Computer Program (Appendix)6. ☒ Nucleotide and/or Amino Acid Sequence Submission

- ☒ Computer Readable Copy
- ☒ Paper Copy (identical to computer copy)
- ☒ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

7. ☐ 37 CFR 3.73(b) Statement (when there is an assignee)
☐ Power of Attorney8. ☐ English Translation Document (if applicable)9. ☐ Information Disclosure Statement (IDS)/PTO-1449
☐ Copies of IDS Citations10. ☐ Preliminary Amendment
☐ Drawings [Total Sheets]
☐ Letter to Official Draftsperson Including Drawings [Total Pages]11. ☒ Return Receipt Postcard12. ☐ Small Entity Statement(s)
☐ Statements filed in prior application, (Status still proper and desired)13. ☐ Certified Copy of Priority Document(s)14. ☐ Deletion of Inventor(s)
Signed statement attached deleting inventor(s) named in the prior application.15. ☒ Patent Application Data Entry Form16. ☐ Other:17. ☐ If a CONTINUING APPLICATION, check appropriate box and supply the requisite information:
☐ Continuation ☐ Divisional ☐ Continuation-in-part (CIP) of prior application Serial No. ____/____.
Priority to the above application(s) is claimed under 35 U.S.C. 120.
Prior application information: Examiner: _____. Group/Art Unit: _____.18. ☒ Priority - 35 U.S.C. 119
☒ Priority of application Serial No. 60/105,507 filed on 10/23/98 in the U.S. and application Serial No. 60/108,685 filed on 11/16/1998 in the U.S. is claimed under 35 U.S.C. 119.
☐ The certified copy has been filed in prior U.S. application Serial No. ____/____ on _____.
☐ The certified copy will follow.

CORRESPONDENCE ADDRESS

Direct all correspondence to: Patent Administrator
Testa, Hurwitz & Thibault, LLP
High Street Tower
125 High Street
Boston, MA 02110
Tel. No.: (617) 248-7000
Fax No.: (617) 248-7100

SIGNATURE BLOCK

Date: 10/22/99
Reg. No. 43,526
Tel. No.: (617) 248-7476
Fax No.: (617) 248-7100Respectfully submitted
Jennifer A. Camacho
Atty/Agent for Applicant(s)
Testa, Hurwitz & Thibault, LLP
High Street Tower
125 High Street
Boston, MA 02110

Inventor Information

Inventor One Given Name :: Hiroaki
 Family Name :: Kawasaki
 Name Suffix ::
 Postal Address Line One :: 3-20-2, Aoba
 Postal Address Line Two ::
 City :: Higashi-Ku
 State/Province :: Fukuoka,
 Country :: JAPAN
 Postal or Zip Code :: 813-0025
 City of Residence :: Higashi-Ku
 State/Prov. of Residence :: Fukuoka,
 Country of Residence :: Japan
 Citizenship :: Japan

Inventor Two Given Name :: Ann
 Family Name :: Graybiel
 Name Suffix ::
 Postal Address Line One :: Boyce Farm Road
 Postal Address Line Two ::
 City :: Lincoln
 State/Province :: Massachusetts
 Country :: U.S.A.
 Postal or Zip Code :: 01773
 City of Residence :: Lincoln
 State/Prov. of Residence :: Massachusetts
 Country of Residence :: U.S.A.
 Citizenship :: U.S.A.

Inventor Three Given Name :: David
 Family Name :: Housman
 Name Suffix ::
 Postal Address Line One :: 64 Homer Street
 Postal Address Line Two ::
 City :: Newton
 State/Province :: Massachusetts
 Country :: U.S.A.
 Postal or Zip Code ::
 City of Residence :: Newton
 State/Prov. of Residence :: Massachusetts
 Country of Residence :: U.S.A.
 Citizenship :: U.S.A.

Correspondence Information

Correspondence Customer Number :: 021323

Application Information

Title Line One :: Genes Integrating Signal Transduction Pathways
 Title Line Two ::
 Total Drawing Sheets :: 3
 Formal Drawings :: N
 Application Type :: Utility
 Docket Number :: MIT-103
 Licensed - U S Government Agency :: NIH
 Contract Numbers: NICHD RO1; NCI PO1; NHLBI PO1; NCHGR RO1
 Grant Numbers:: HD28341; CA42063; HL41484; HG00299
 Secrecy Order in Parent Application ::

Representative Information

Representative Customer Number :: 021323

Continuity Information

This application:: Claims Benefit of
 >Application One :: 60/105,507
 Filing Date :: October 23, 1998
 Patent Number ::

This application:: Claims Benefit of
 >Application Two :: 60/108,685
 Filing Date :: November 16, 1998
 Patent Number ::

Prior Foreign Applications

Foreign Application One ::
 Filing Date ::
 Country ::
 Priority Claimed ::

479JAC5473/111.A877130-1

GENES INTEGRATING SIGNAL TRANSDUCTION PATHWAYSRelated Applications

This application claims the benefit of U.S. Application Nos. 60/105,507, filed on October 23, 1998, and 60/108,685, filed on November 16, 1998.

5 Field of the Invention

The present invention relates generally to novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic, and research utilities for these polynucleotides and proteins.

10 Background of the Invention

Ras proteins are key regulators of growth, differentiation and malignant transformation. In addition, these proteins are implicated in synaptic function and region-specific learning and memory functions in the brain.

As shown schematically in Figure 1, Ras proteins cycle between inactive GDP-complexed and active GTP-complexed states. GTPase-activating proteins (GAPs) inactivate Ras proteins by stimulating hydrolysis of the bound GTP to GDP, whereas guanine nucleotide exchange factors (GEFs) activate Ras proteins by stimulating release of GDP and the uptake of GTP. So essential are GEFs to Ras action, that genetic loss of GEF function has similar effects to those induced by loss of the Ras proteins themselves. Loss of GEF function can be
15 circumvented by mutations that constitutively activate the Ras proteins, such as an oncogene mutation, or, in some cases, through loss of GAP activity. Activated Ras proteins, which are localized at the plasma membrane, transmit signals from tyrosine kinases to a cascade of serine/threonine kinases, which delivers the signals to the cell nucleus.

Activation of Ras can result in the activation of the mitogen-activated protein (MAP)
25 kinase (also known as extracellular-signal regulated kinase, or ERK) pathway. For example, a receptor tyrosine kinase is activated by a peptide mitogen such as epidermal growth factor

(EGF). The EGF-stimulated receptor undergoes autophosphorylation of specific tyrosine residues in its cytoplasmic domain which creates phosphotyrosyl binding sites for the Src homology 2 (SH2) and/or phosphotyrosyl binding (PTB) domains of certain adapter proteins. The adapter protein becomes autophosphorylated on association with activated receptor tyrosine
5 kinases. The GEF is stably associated with the adapter protein which, upon autophosphorylation, mediates translocation of the GEF to the plasma membrane. The GEF then activates the Ras protein. Activated Ras relays its signal downstream through a cascade of cytoplasmic proteins, including Raf-1 serine/threonine kinase. The Ras:Raf association promotes translocation of the normally cytoplasmic Raf protein to the plasma membrane, where subsequent events lead to the
10 activation of its kinase function. Upon activation, Raf phosphorylates and activates two MAP kinases (also known as MEKs). MEKs directly associate with the catalytic domain of Raf-1 and are phosphorylated by Raf. Activated MEKs function as dual-specificity kinases and phosphorylate tandem threonine and tyrosine residues in the MAP kinases to activate them. Once activated, the MAP kinases translocate to the nucleus where they phosphorylate and
15 activate a variety of substrates.

Rap proteins, members of the Ras small GTPase superfamily, can inhibit Ras signaling of the Ras/Raf-1(a serine/threonine kinase)/MAP kinase pathway or, through B-Raf, can activate MAP kinase. Rap1 consists of two isoforms, Rap1A and Rap1B, which differ mainly at the C-terminus. Characteristic features of Rap1 are its geranylgeranyl modification at the C-terminus,
20 which is responsible for membrane attachments, and a threonine residue at position 61. In most other GTPases, the corresponding residue is a glutamine. Rap proteins, like Ras proteins, cycle between inactive GDP-complexed and active GTP-complexed states. GEFs are required to activate Rap proteins by stimulating the release of GDP and the uptake of GTP.

Constitutive activation of the Ras pathway contributes to malignant transformation. In
25 fact, the Ras gene has been implicated in many human cancers, including lung cancer, breast cancer, colorectal cancer, exocrine pancreatic cancer, and myeloid leukemia. Biological and biochemical studies of Ras action indicate that Ras functions like a G-regulatory protein since Ras must be localized in the plasma membrane and must bind with GTP in order to transform cells. Gibbs et al., 53 MICROBIOL. REV. 171-286 (1989).

Targeting components of the Ras signaling pathways has been proposed as one approach for the development of anti-Ras drugs for cancer treatment. One potential approach for targeting Ras for cancer treatment involves the use of farnesyltransferase inhibitors (FTIs). Inhibition of farnesyl-protein transferase, and thereby of farnesylation of the Ras protein, blocks the ability of Ras to transform normal cells to cancer cells. Certain inhibitors of Ras farnesylation cause an increase in soluble Ras which can act as a dominant negative inhibitor of Ras function. While soluble Ras in cancer cells can become a dominant negative inhibitor, soluble Ras in normal cells would not be an inhibitor. A cytosol-localized and activated form of Ras acts as a dominant negative Ras inhibitor of membrane-bound Ras function. Gibbs et al., 86 PROC. NAT'L ACAD. SCI. USA 6630-34 (1989). FTIs block Ras function by preventing its post-translational modification by the farnesyl isoprenoid.

Intervention of Ras signaling at multiple or various points can significantly impact the ability of Ras to cause cellular transformation. Since Ras protein function is believed to be crucial to so many cellular processes, targeting only a subset of Ras functions by downstream intervention may provide significant advantages. Thus, there remains a need for identifying additional means for disrupting the Ras pathway. Applicants have discovered four new targets, namely GEFs specific for Rap1A , for disrupting the Ras pathway.

Summary of the Invention

Applicants have discovered four mammalian genes which have been designated CalDAG-GEFI, CalDAG-GEFII, cAMP-GEFI, and cAMP-GEFII, which encode proteins having a substrate specificity for Rap1A. The proteins encoded by CalDAG-GEFI and CalDAG-GEFII, referred to herein generally as "CalDAG-GEF," have dual binding domains for calcium and diacylglycerol. The proteins cAMP-GEFI and cAMP-GEFII, referred to herein generally as "cAMP-GEF," have a binding domain for cyclic adenosine 3', 5'-monophosphate. The present disclosure provides polypeptide and polynucleotide sequences for *Mus musculus* CalDAG-GEFI, *Homo sapiens* CalDAG-GEFI, *Rattus norvegicus* CalDAG-GEFII, *Homo sapiens* CalDAG-GEFII, *Rattus norvegicus* cAMP-GEFI, *Homo sapiens* cAMP-GEFI, *Homo sapiens* alternatively spliced cAMP-GEFI, *Rattus norvegicus* cAMP-GEFII, and *Homo sapiens* cAMP-GEFII. See

Kawasaki et al., 95 Proc. Natl. Acad. Sci. USA 13278-83 (1998), and Kawasaki et al., 282 Sci. 2275-79 (1998), the disclosures of both of which are incorporated by reference herein.

Thus, in one series of embodiments, the present invention provides isolated nucleic acids including nucleotide sequences comprising or derived from CalDAG-GEF or cAMP-GEF, or encoding polypeptides comprising or derived from CalDAG-GEF or cAMP-GEF proteins. The sequences of the invention include the specifically disclosed sequences, splice variants of these sequences, allelic variants of these sequences, synonymous sequences, and homologous or orthologous variants of these sequences. Thus, for example, the invention provides nucleic acid sequences from the *Mus musculus* CalDAG-GEFI, *Homo sapiens* CalDAG-GEFI, *Rattus norvegicus* CalDAG-GEFII, *Homo sapiens* CalDAG-GEFII, *Rattus norvegicus* cAMP-GEFI, *Homo sapiens* cAMP-GEFI, *Homo sapiens* alternatively spliced cAMP-GEFI, *Rattus norvegicus* cAMP-GEFII, and *Homo sapiens* cAMP-GEFII. The present invention also provides allelic variants and homologous or orthologous sequences by providing methods by which such variants may be routinely obtained. Because the nucleic acids of the invention may be used in a variety of diagnostic, therapeutic and recombinant applications, various subsets of the CalDAG-GEF and cAMP-GEF sequences are also provided. For example, for use in allele specific hybridization screening or PCR amplification techniques, subsets of the CalDAG-GEF and cAMP-GEF sequences, including both sense and antisense sequences, and both normal and mutant sequences, as well as intronic, exonic and untranslated sequences, are provided. Such sequences may comprise a small number of consecutive nucleotides from the sequences which are disclosed or otherwise enabled herein, but preferably include at least 8-10, more preferably 10-15, and most preferably 15-25, consecutive nucleotides from a CalDAG-GEF or cAMP-GEF sequence. In another embodiment, such sequences include between 25-500 consecutive nucleotides. Other preferred subsets of a CalDAG-GEF or cAMP-GEF sequence include those encoding one or more of the functional domains or antigenic determinants of the CalDAG-GEF or cAMP-GEF protein and, in particular, may include either normal (wild-type) or mutant sequences. The invention also provides for various nucleic acid constructs in which CalDAG-GEF or cAMP-GEF sequences, either complete or subsets, are operably joined to exogenous sequences to form cloning vectors, expression vectors, fusion vectors, transgenic constructs, and the like. Thus, in accordance with another aspect of the invention, a recombinant vector for transforming a

mammalian or invertebrate tissue cell to express a normal or mutant CalDAG-GEF and/or cAMP-GEF sequence in the cells is provided.

In another series of embodiments, the present invention provides for host cells which have been transfected or otherwise transformed with one of the nucleic acids of the invention.

5 The cells may be transformed merely for purposes of propagating the nucleic acid constructs of the invention, or may be transformed so as to express the CalDAG-GEF and/or cAMP-GEF sequences. The transformed cells of the invention may be used in assays to identify proteins and/or other compounds which affect normal or mutant CalDAG-GEF and/or cAMP-GEF expression, which interact with the normal or mutant CalDAG-GEF and/or cAMP-GEF proteins,
10 and/or which modulate the function or effects of the normal or mutant proteins, or to produce the CalDAG-GEF and/or cAMP-GEF proteins, fusion proteins, functional domains, antigenic determinants, and/or antibodies of the invention. Transformed cells may also be implanted into hosts, including humans, for therapeutic or other reasons. Preferred host cells include mammalian cells, including pure or mixed cell cultures, as well as bacterial, yeast, nematode,
15 insect and other invertebrate cells. For uses as described below, preferred cells also include embryonic stem cells, zygotes, gametes, and germ line cells.

In another series of embodiments, the present invention provides transgenic animal models of diseases or disorders associated with mutations in the CalDAG-GEF and/or cAMP-GEF genes. The animal may be essentially any non-human mammal, including rats, mice,
20 hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates. In addition, invertebrate models, including nematodes and insects, may be used for certain applications. The animal models are produced by standard transgenic methods including microinjection, electroporation, transfection, or other forms of transformation of embryonic stem cells, zygotes, gametes, and germ line cells with vectors including genomic or cDNA fragments,
25 minigenes, homologous recombination vectors, viral insertion vectors and the like. Suitable vectors include vaccinia virus, adenovirus, adeno-associated virus, retrovirus, liposome transport, neurotropic viruses, and Herpes simplex virus. The animal models may include transgenic sequences comprising or derived from the CalDAG-GEF and/or cAMP-GEF genes, including normal and mutant sequences, intronic, exonic and untranslated sequences, and sequences
30 encoding subsets of the CalDAG-GEF and/or cAMP-GEF proteins, such as functional domains.

The major types of animal models provided include: (1) Animals in which a normal human CalDAG-GEF and/or cAMP-GEF gene has been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a recombinant gene or a large genomic fragment; in which a

5 normal human CalDAG-GEF and/or cAMP-GEF gene has been recombinantly substituted for one or both copies of the animal's homologous CalDAG-GEF and/or cAMP-GEF gene by homologous recombination or gene targeting; and/or in which one or both copies of one of the animal's homologous CalDAG-GEF and/or cAMP-GEF genes have been recombinantly

10 "humanized" by the partial substitution of sequences encoding the human homologue by homologous recombination or gene targeting; (2) Animals in which a mutant human CalDAG-GEF and/or cAMP-GEF gene has been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a recombinant gene or a large genomic fragment; in which a mutant human

15 CalDAG-GEF and/or cAMP-GEF gene has been recombinantly substituted for one or both copies of the animal's homologous CalDAG-GEF and/or cAMP-GEF gene by homologous recombination or gene targeting; and/or in which one or both copies of one of the animal's homologous CalDAG-GEF and/or cAMP-GEF gene have been recombinantly "humanized" by the partial substitution of sequences encoding a mutant human homologue by homologous recombination or gene targeting; (3) Animals in which a mutant version of one of that animal's

20 CalDAG-GEF or cAMP-GEF gene has been recombinantly introduced into the genome of the animal as an additional gene, under the regulation of either an exogenous or an endogenous promoter element, and as either a recombinant gene or a large genomic fragment; and/or in which a mutant version of one of that animal's CalDAG-GEF or cAMP-GEF gene has been recombinantly substituted for one or both copies of the animal's homologous CalDAG-GEF or

25 cAMP-GEF gene by homologous recombination or gene targeting; and (4) "Knock-out" animals in which one or both copies of one of the animal's CalDAG-GEF or cAMP-GEF genes have been partially or completely deleted by homologous recombination or gene targeting, or have been inactivated by the insertion or substitution by homologous recombination or gene targeting of exogenous sequences.

In another series of embodiments, the present invention provides for substantially pure protein preparations including polypeptides comprising or derived from the CalDAG-GEF and/or cAMP-GEF proteins. The CalDAG-GEF and cAMP-GEF protein sequences of the invention include the specifically disclosed sequences, variants of these sequences resulting from

5 alternative mRNA splicing, allelic variants of these sequences, and homologous or orthologous variants of these sequences. Thus, for example, the invention provides amino acid sequences from the *Mus musculus* CalDAG-GEFI protein, *Homo sapiens* CalDAG-GEFI protein, *Rattus norvegicus* CalDAG-GEFII protein, *Homo sapiens* CalDAG-GEFII protein, *Rattus norvegicus* cAMP-GEFI protein, *Homo sapiens* cAMP-GEFI protein, *Homo sapiens* alternatively spliced
10 cAMP-GEFI protein, *Rattus norvegicus* cAMP-GEFII protein, and *Homo sapiens* cAMP-GEFII protein. The present invention also provides allelic variants and homologous or orthologous proteins by providing methods by which such variants may be routinely obtained. The present invention also specifically provides for mutant or disease-causing variants of CalDAG-GEF and cAMP-GEF by providing methods by which such variants may be routinely obtained. Because
15 the proteins of the invention may be used in a variety of diagnostic, therapeutic and recombinant applications, various subsets of the CalDAG-GEF and cAMP-GEF protein sequences and combinations of the CalDAG-GEF and cAMP-GEF protein sequences with heterologous sequences are also provided. For example, for use as immunogens or in binding assays, subsets of the CalDAG-GEF and cAMP-GEF protein sequences, including both normal and mutant
20 sequences, are provided. Such protein sequences may comprise a small number of consecutive amino acid residues from the sequences which are disclosed or otherwise enabled herein, but preferably include at least 4-8, and preferably at least 9-15 consecutive amino acid residues from a CalDAG-GEF or cAMP-GEF sequence. In another embodiment, such sequences comprise at least 15-200 consecutive amino acid residues. Other preferred subsets of the CalDAG-GEF and
25 cAMP-GEF protein sequences include those corresponding to one or more of the functional domains or antigenic determinants of the CalDAG-GEF and cAMP-GEF proteins and, in particular, may include either normal (wild-type) or mutant sequences. The invention also provides for various protein constructs in which a CalDAG-GEF and/or cAMP-GEF sequences, either complete or subsets thereof, are joined to exogenous sequences to form fusion proteins and
30 the like. In accordance with these embodiments, the present invention also provides for methods

of producing all of the above described proteins which comprise, or are derived from, CalDAG-GEF and/or cAMP-GEF.

In another series of embodiments, the present invention provides for the production and use of polyclonal and monoclonal antibodies, including antibody fragments, including Fab fragments, $F(ab')_2$, and single chain antibody fragments, which selectively bind to CalDAG-GEF or cAMP-GEF, or to specific antigenic determinants of CalDAG-GEF or cAMP-GEF. The antibodies may be raised in mouse, rabbit, goat or other suitable animals, or may be produced recombinantly in cultured cells such as hybridoma cell lines. Preferably, the antibodies selectively bind to a sequence comprising at least 4-8, and preferably at least 9-15, and more preferably at least 15-200 consecutive amino acid residues from a CalDAG-GEF or cAMP-GEF sequence. The antibodies of the invention may be used in the various diagnostic, therapeutic and technical applications described herein.

In another series of embodiments, the present invention provides methods of screening or identifying proteins, small molecules or other compounds which are capable of inducing or inhibiting the expression and/or function of the CalDAG-GEF and/or cAMP-GEF genes or proteins. The assays may be performed *in vitro* using non-transformed cells, immortalized cell lines, or recombinant cell lines, or *in vivo* using the transgenic animal models enabled herein. In particular, the assays may detect the presence of increased or decreased expression of CalDAG-GEF and/or cAMP-GEF-related genes or proteins on the basis of increased or decreased mRNA expression, increased or decreased levels of CalDAG-GEF and/or cAMP-GEF-related protein products, or increased or decreased levels of expression of a marker gene (*e.g.*, β -galactosidase, green fluorescent protein, alkaline phosphatase or luciferase) operably joined to a 5' regulatory region in a recombinant construct. Cells known to express CalDAG-GEF or cAMP-GEF, or transformed to express CalDAG-GEF or cAMP-GEF, are incubated and one or more test compounds are added to the medium. After allowing a sufficient period of time (*e.g.*, 0-72 hours) for the compound to induce or inhibit the expression of the CalDAG-GEF or cAMP-GEF, any change in levels of expression from an established baseline may be detected using any of the techniques described above. In particularly preferred embodiments, the cells are from an immortalized cell line such as a human neuroblastoma, glioblastoma or a hybridoma cell line, or are transformed cells of the invention.

In another series of embodiments, the present invention provides methods for identifying proteins and other compounds which bind to, or otherwise directly interact with, CalDAG-GEF and/or cAMP-GEF. The proteins and compounds will include endogenous cellular components which interact with the CalDAG-GEF and/or cAMP-GEF *in vivo* and which, therefore, provide new targets for pharmaceutical and therapeutic interventions, as well as recombinant, synthetic, and otherwise exogenous compounds which may have CalDAG-GEF and/or cAMP-GEF binding capacity and, therefore, may be candidates for pharmaceutical agents. Particularly useful components will bind with a SCR1, SCR2, SCR3, EF hand, DAG-binding or cAMP-binding domain. Thus, in one series of embodiments, cell lysates or tissue homogenates (*e.g.*, human brain homogenates, lymphocyte lysates) may be screened for proteins or other compounds which bind to one of the normal or mutant CalDAG-GEF or cAMP-GEF proteins. Alternatively, any of a variety of exogenous compounds, both naturally occurring and/or synthetic (*e.g.*, libraries of small molecules or peptides), may be screened for CalDAG-GEF or cAMP-GEF binding capacity. In each of these embodiments, an assay is conducted to detect binding between a "CalDAG-GEF component" or a "cAMP-GEF component" and some other moiety. The "CalDAG-GEF component" or the "cAMP-GEF component" in these assays may be any polypeptide comprising or derived from a normal or mutant CalDAG-GEF or cAMP-GEF protein, including functional domains or antigenic determinants of CalDAG-GEF or cAMP-GEF, or CalDAG-GEF or cAMP-GEF fusion proteins. In one embodiment, a CalDAG-GEF component is a SCR1, SCR2, SCR3, EF hand, or DAG-binding domain. In another embodiment a cAMP-GEF component is a SCR1, SCR2, SCR3, or cAMP-binding domain. Binding may be detected by non-specific measures (*e.g.*, changes in intracellular Ca^{2+} , GTP/GDP ratio) or by specific measures (*e.g.*, changes in the expression of downstream genes which can be monitored by differential display, 2D gel electrophoresis, differential hybridization, or SAGE methods). The preferred methods involve variations on the following techniques: (1) direct extraction by affinity chromatography; (2) co-isolation of CalDAG-GEF or cAMP-GEF components and bound proteins or other compounds by immunoprecipitation; (3) the Biomolecular Interaction Assay (BIAcore); and (4) the yeast two-hybrid systems.

In another series of embodiments, the present invention provides for methods of identifying proteins, small molecules and other compounds capable of modulating the activity of

normal or mutant CalDAG-GEF or cAMP-GEF. Using normal cells or animals, the transformed cells and transgenic animal models of the present invention, or cells obtained from subjects bearing normal or mutant CalDAG-GEF or cAMP-GEF genes, the present invention provides methods of identifying such compounds on the basis of their ability to affect the expression of CalDAG-GEF and/or cAMP-GEF, the intracellular localization of the CalDAG-GEF and/or cAMP-GEF, or other biochemical, histological, or physiological markers which distinguish cells bearing normal and mutant CalDAG-GEF and/or cAMP-GEF sequences. Using the transgenic animals of the invention, methods of identifying such compounds are also provided on the basis of the ability of the compounds to affect behavioral, physiological or histological phenotypes associated with mutations in CalDAG-GEF and/or cAMP-GEF.

In another series of embodiments, the present invention provides methods and reagents for the screening and diagnosis of diseases or disorders associated with mutations in the CalDAG-GEF and/or cAMP-GEF genes. Screening and/or diagnosis can be accomplished by methods based upon the nucleic acids (including genomic and mRNA/cDNA sequences), proteins, and/or antibodies disclosed and enabled herein, including functional assays designed to detect failure or augmentation of the normal CalDAG-GEF and/or cAMP-GEF activity and/or the presence of specific new activities conferred by the mutant CalDAG-GEF and/or cAMP-GEF. Thus, for example, screens and diagnostics based upon CalDAG-GEF and/or cAMP-GEF proteins are provided which detect differences between mutant and normal CalDAG-GEF or cAMP-GEF in electrophoretic mobility, in proteolytic cleavage patterns, in molar ratios of the various amino acid residues, or in ability to bind specific antibodies. In addition, screens and diagnostics based upon nucleic acids (gDNA, cDNA or mRNA) are provided which detect differences in nucleotide sequences by direct nucleotide sequencing, hybridization using allele specific oligonucleotides, restriction enzyme digest and mapping (*e.g.*, RFLP, REF-SSCP), electrophoretic mobility (*e.g.*, SSCP, DGGE), PCR mapping, RNase protection, chemical mismatch cleavage, ligase-mediated detection, and various other methods. Other methods are also provided which detect abnormal processing of CalDAG-GEF and/or cAMP-GEF or proteins reacting with CalDAG-GEF and/or cAMP-GEF, alterations in CalDAG-GEF and/or cAMP-GEF transcription, translation, and post-translational modification; alterations in the intracellular and extracellular trafficking of CalDAG-GEF and/or cAMP-GEF gene products; or abnormal

intracellular localization of CalDAG-GEF and/or cAMP-GEF. Such methods and reagents are also useful in the analysis of neoplasias and mammalian immune system function, as well as functional *in vivo* imaging of mammalian organ systems. In accordance with these embodiments, diagnostic kits are also provided which will include the reagents necessary for the

5 above-described diagnostic screens.

In another series of embodiments, the present invention provides methods and therapeutic agents for use in the treatment of conditions such as neurological and neuropsychiatric disorders such as Huntington's disease, Parkinson's disease, Alzheimer's disease, dystonia, Tourette's syndrome, obsessive-compulsive disorder, attention deficit/hyperactive disorder, depression,

10 schizophrenia, and stroke; neoplasias such as solid tumors including colon, breast, lung, prostate, and hematopoietic tumors such as leukemia, Hodgkins, and non-Hodgkins lymphomas; and autoimmune diseases, allergies, and asthma; as well as for the enhancement of the immune response in normal and immunocompromised individuals. These methods and therapeutic agents may be based upon (1) administration of normal CalDAG-GEF and/or cAMP-GEF proteins; (2)

15 gene therapy with normal CalDAG-GEF and/or cAMP-GEF genes to compensate for or replace the mutant genes; (3) gene therapy based upon antisense sequences to mutant CalDAG-GEF and/or cAMP-GEF genes or upon sequences which "knock-out" the mutant genes; (4) gene therapy based upon sequences which encode a protein which blocks or corrects the deleterious effects of CalDAG-GEF and/or cAMP-GEF mutants; (5) immunotherapy based upon antibodies

20 to normal and/or mutant CalDAG-GEF and/or cAMP-GEF proteins; or (6) small molecules (drugs) which alter CalDAG-GEF and/or cAMP-GEF expression, block interactions between (normal or mutant) forms of CalDAG-GEF and/or cAMP-GEF and other proteins or ligands, or which otherwise block the function of (normal or mutant) CalDAG-GEF and/or cAMP-GEF proteins by altering the structure of the proteins, by enhancing their metabolic clearance, or by

25 inhibiting their function.

In accordance with another aspect of the invention, the proteins of the invention can be used as starting points for rational drug design to provide ligands, therapeutic drugs or other types of small chemical molecules. Alternatively, small molecules or other compounds identified by the above-described screening assays may serve as "lead compounds" in rational

30 drug design.

Brief Description of the Drawings

Figure 1 is a partial schematic diagram of a Ras pathway.

Figure 2A shows human (h) and mouse (m) CalDAG-GEFI, human (h) and rat (r)

5 CalDAG-GEFII, and *C. elegans* (cel) (F25B3.3, GenBank accession number: 1262950) CalDAG-GEF. Figure 2B shows a computer-generated phylogenetic tree analysis of the GEF domains of CalDAG-GEFI and CalDAG-GEFII in relation to other Ras-superfamily GEFs. Figure 2C shows multiple alignment of GEF structurally conserved regions (SCRs) of CalDAG-GEFs and several other GEFs of the Ras superfamily. Figure 2D shows the full-length amino
10 acid sequences of human (h) and mouse (m) CalDAG-GEFI (box indicates amino acid differences). Figure 2E shows the sequence similarity (black indicates identity) of EF-hand domains in CalDAG-GEFs and other calcium binding proteins. Figure 2F shows the sequence similarity of DAG-binding domains of CalDAG-GEFs and PKC (protein kinase C) family proteins.

15 Figure 3A is a schematic representation of cAMP-GEF family proteins, including human (h) and rat (r) cAMP-GEFI, human (h) cAMP-GEFII and *C. elegans* (cel) (T2OG5.5, GenBank accession number: 458480) cAMP-GEF. Figure 3B is a phylogenetic tree analysis of cAMP binding domains of cAMP-GEFI and II and other cyclic nucleotide binding proteins. Figure 3C is a phylogenetic tree analysis of GEF domains of cAMP-GEFI and II and other Ras superfamily
20 GEFs. Figure 3D shows the amino acid sequences of the structurally conserved regions (SCRs) of cAMP-GEFs and other Ras superfamily GEFs (black indicates identity). Figure 3E shows the amino acid sequences of the cAMP binding pockets of cAMP-GEFI and II and other cyclic nucleotide-binding proteins. The positions of invariant amino acid residues are shown by black diamonds. The open diamond indicates the amino acid that determines the binding specificity
25 for cAMP or cGMP. The arrow indicates the position of amino acid substitutions specific to cAMP-GEFs. Figure 3F is the full-length amino acid sequences of human cAMP-GEFI and II (boxes indicate amino acid identity).

Detailed Description of the Invention

The present invention is based, in part, upon the discovery of a family of mammalian genes which are associated with the Ras pathway. The discovery of these genes, designated CalDAG-GEFI, CalDAG-GEFII, cAMP-GEFI, and cAMP-GEFII, as well as the characterization of these genes, their protein products, mutants, and possible functional roles, are described below.

I. Definitions

In order to facilitate review of the various embodiments of the invention, and an understanding of the various elements and constituents used in making and using the invention, the following definitions are provided for particular terms used in the description and the claims which follow:

CalDAG-GEF. As used without further modification herein, the terms "CalDAG-GEF" or "CalDAG-GEFs" refer to the CalDAG-GEFI and/or the CalDAG-GEFII genes/proteins. In particular, the unmodified terms "CalDAG-GEF" or "CalDAG-GEFs" refer to the mammalian genes/proteins and, preferably, the human genes/proteins.

cAMP-GEF. As used without further modification herein, the terms "cAMP-GEF" or "cAMP-GEFs" refer to the cAMP-GEFI and/or the cAMP-GEFII genes/proteins. In particular, the unmodified terms "cAMP-GEF" or "cAMP-GEFs" refer to the mammalian genes/proteins and, preferably, the human genes/proteins.

CalDAG-GEF gene. As used herein, the term "CalDAG-GEF gene" means the mammalian genes represented by SEQ ID NOS: 1, 3, 5, and 7, as well as any allelic variants and heterospecific mammalian homologues. A murine CalDAG-GEFI cDNA sequence is disclosed herein as SEQ ID NO: 1, and a human CalDAG-GEFI cDNA sequence is disclosed herein as SEQ ID NO: 3. A rat CalDAG-GEFII cDNA sequence is disclosed herein as SEQ ID NO: 5, and a human CalDAG-GEFII cDNA sequence is disclosed herein as SEQ ID NO: 7. The term "CalDAG-GEF gene" primarily relates to a coding sequence, but can also include some or all of the flanking regulatory regions and/or introns. The term "CalDAG-GEF gene" specifically includes artificial or recombinant genes created from cDNA or genomic DNA, including recombinant genes based upon splice variants.

CalDAG-GEF protein. As used herein, the term "CalDAG-GEF protein" means a protein encoded by a CalDAG-GEF gene, including allelic variants and heterospecific mammalian homologues. A murine CalDAG-GEFI protein sequence is disclosed herein as SEQ ID NO: 2, and a human CalDAG-GEFI protein sequence is disclosed herein as SEQ ID NO: 4. A rat

5 CalDAG-GEFII protein sequence is disclosed herein as SEQ ID NO: 6, and a human CalDAG-GEFII protein sequence is disclosed herein as SEQ ID NO: 8. Splice variants are also embraced by the term CalDAG-GEF protein as used herein. The protein may be produced by recombinant cells or organisms, may be substantially purified from natural tissues or cell lines, or may be synthesized chemically or enzymatically. Therefore, the term "CalDAG-GEF protein" is

10 intended to include the protein in glycosylated, partially glycosylated, or unglycosylated forms, as well as in phosphorylated, partially phosphorylated, unphosphorylated, sulphated, partially sulphated, or unsulphated forms. The term also includes allelic variants and other functional equivalents of the CalDAG-GEF amino acid sequences, including biologically active proteolytic or other fragments.

15 hCalDAG-GEF gene and/or protein. As used herein, the abbreviation "hCalDAG-GEF" refers to the human homologue and human allelic variants of the CalDAG-GEF genes and/or proteins. Two cDNA sequences of the human CalDAG-GEF genes are disclosed herein as SEQ ID NOS: 3 and 7. The corresponding hCalDAG-GEF protein sequences are disclosed herein as SEQ ID NOS: 4 and 8. Allelic variants, including deleterious mutants, are enabled in the

20 description which follows.

mCalDAG-GEF gene and/or protein. As used herein, the abbreviation "mCalDAG-GEF" refers to the murine homologues and murine allelic variants of the CalDAG-GEF gene and/or protein. A cDNA sequence of one murine CalDAG-GEF gene is disclosed herein as SEQ ID NO: 16. The corresponding mCalDAG-GEF protein sequence is disclosed herein as SEQ ID

25 NO: 17. Allelic variants, including deleterious mutants, are enabled in the description which follows.

rCalDAG-GEF gene and/or protein. As used herein, the abbreviation "rCalDAG-GEF" refers to the rat homologue and rat allelic variants of the CalDAG-GEF genes and/or proteins. A cDNA sequence of one rat CalDAG-GEF gene is disclosed herein as SEQ ID NO: 5. The

corresponding rCalDAG-GEF protein sequence is disclosed herein as SEQ ID NO: 6. Allelic variants, including deleterious mutants, are enabled in the description which follows.

cAMP-GEF gene. As used herein, the term "cAMP-GEF gene" means the mammalian genes represented by SEQ ID NOS: 9, 11, 13, 15, and 17, as well as any allelic variants and heterospecific mammalian homologues. A rat cAMP-GEFI cDNA sequence is disclosed herein as SEQ ID NO: 9, and a human cAMP-GEFI cDNA sequence is disclosed as SEQ ID NO:11. Another human cAMP-GEFI cDNA sequence, resulting from alternative splicing of the mRNA transcript, is disclosed as SEQ ID NO: 13. A rat cAMP-GEFII cDNA sequence is disclosed as SEQ ID NO: 15, and a human cAMP-GEFII cDNA sequence is disclosed as SEQ ID NO: 17.

The term "cAMP-GEF gene" primarily relates to a coding sequence, but can also include some or all of the flanking regulatory regions and/or introns. The term cAMP-GEF gene specifically includes artificial or recombinant genes created from cDNA or genomic DNA, including recombinant genes based upon splice variants.

cAMP-GEF protein. As used herein, the term "cAMP-GEF protein" means a protein encoded by a cAMP-GEF gene, including allelic variants and heterospecific mammalian homologues. A rat cAMP-GEFI protein sequence is disclosed herein as SEQ ID NO: 10, and a human cAMP-GEFI protein sequence is disclosed as SEQ ID NO: 12. Another human cAMP-GEFI protein sequence, resulting from alternative splicing of the mRNA transcript, is disclosed as SEQ ID NO: 14. A rat cAMP-GEFII protein sequence is disclosed as SEQ ID NO: 16, and a human cAMP-GEFII protein sequence is disclosed as SEQ ID NO: 18. Splice variants are also embraced by the term cAMP-GEF protein as used herein. The protein may be produced by recombinant cells or organisms, may be substantially purified from natural tissues or cell lines, or may be synthesized chemically or enzymatically. Therefore, the term "cAMP-GEF protein" is intended to include the protein in glycosylated, partially glycosylated, or unglycosylated forms, as well as in phosphorylated, partially phosphorylated, unphosphorylated, sulphated, partially sulphated, or unsulphated forms. The term also includes allelic variants and other functional equivalents of the cAMP-GEF amino acid sequences, including biologically active proteolytic or other fragments.

hcAMP-GEF gene and/or protein. As used herein, the abbreviation "hcAMP-GEF" refers to the human homologue and human allelic variants of the cAMP-GEF gene and/or protein. One

cDNA sequences of the human cAMP-GEF gene is disclosed herein as SEQ ID NO: 18. The corresponding hcAMP-GEF protein sequence is disclosed herein as SEQ ID NO: 19. Numerous allelic variants, including deleterious mutants, are disclosed and enabled throughout the description which follows.

5 rcAMP-GEF gene and/or protein. As used herein, the abbreviation "rcAMP-GEF" refers to the rat homologue and rat allelic variants of the cAMP-GEF gene and/or protein. Two cDNA sequences of rat cAMP-GEF genes are disclosed herein as SEQ ID NOS: 9 and 15. The corresponding rcAMP-GEF protein sequences are disclosed herein as SEQ ID NOS: 10 and 16.. Numerous allelic variants, including deleterious mutants, are disclosed and enabled throughout
10 the description which follows.

Normal. As used herein with respect to genes, the term "normal" refers to a gene which encodes and expresses a normal protein. As used herein with respect to proteins, the term "normal" means a protein which performs its usual or normal physiological role and which is not associated with, or causative of, a pathogenic condition or state. Therefore, as used herein,
15 the term "normal" is essentially synonymous with the usual meaning of the phrase "wild type." For any given gene, or corresponding protein, a multiplicity of normal allelic variants may exist, none of which is associated with the development of a pathogenic condition or state. Such normal allelic variants include, but are not limited to, variants in which one or more nucleotide substitutions do not result in a change in the encoded amino acid sequence.

20 Mutant. As used herein with respect to genes, the term "mutant" refers to a gene which encodes a mutant protein and/or fails to express a normal protein. As used herein with respect to proteins, the term "mutant" means a protein which does not perform its usual or normal physiological role and which is associated with, or causative of, a pathogenic condition or state. Therefore, as used herein, the term "mutant" is essentially synonymous with the terms
25 "dysfunctional," "pathogenic," "disease-causing," and "deleterious." With respect to the CalDAG-GEF and cAMP-GEF genes and proteins of the present invention, the term "mutant" refers to CalDAG-GEF and cAMP-GEF genes/proteins bearing one or more nucleotide/amino acid substitutions, insertions and/or deletions which cause the genes/proteins to be dysfunctional, pathogenic, disease-causing or otherwise deleterious. This definition is understood to include the
30 various mutations that naturally exist, including but not limited to those disclosed herein, as well

as synthetic or recombinant mutations produced by human intervention. The term "mutant," as applied to the CalDAG-GEF and cAMP-GEF genes, is not intended to embrace sequence variants which, due to the degeneracy of the genetic code, encode proteins identical to the normal sequences disclosed or otherwise enabled herein; nor is it intended to embrace sequence variants which, although they encode different proteins, encode proteins which are functionally equivalent to normal CalDAG-GEF and/or cAMP-GEF proteins.

Functional equivalent. As used herein in describing gene sequences and amino acid sequences, the term "functional equivalent" means that a recited sequence need not be identical to a particularly disclosed sequence of the SEQ ID NOs but need only provide a sequence which functions biologically and/or chemically as the equivalent of the disclosed sequence.

Substantially pure. As used herein with respect to protein preparations, the term "substantially pure" means a preparation which contains at least 60% (by dry weight) the protein of interest, exclusive of the weight of other intentionally included compounds. Preferably the preparation is at least 75%, more preferably at least 90%, and most preferably at least 99%, by dry weight the protein of interest, exclusive of the weight of other intentionally included compounds. Purity can be measured by any appropriate method, e.g., column chromatography, gel electrophoresis, or HPLC analysis. If a preparation intentionally includes two or more different proteins of the invention, a "substantially pure" preparation means a preparation in which the total dry weight of the proteins of the invention is at least 60% of the total dry weight, exclusive of the weight of other intentionally included compounds. Preferably, for such preparations containing two or more proteins of the invention, the total weight of the proteins of the invention be at least 75%, more preferably at least 90%, and most preferably at least 99%, of the total dry weight of the preparation, exclusive of the weight of other intentionally included compounds. Thus, if the proteins of the invention are mixed with one or more other proteins (e.g., serum albumin, 6-OST) or compounds (e.g., diluents, detergents, excipients, salts, polysaccharides, sugars, lipids) for purposes of administration, stability, storage, and the like, the weight of such other proteins or compounds is ignored in the calculation of the purity of the preparation.

Isolated nucleic acid. As used herein, an "isolated nucleic acid" is a ribonucleic acid, deoxyribonucleic acid, or nucleic acid analog comprising a polynucleotide sequence that has

been isolated or separated from sequences that are immediately contiguous (one on the 5' end and one on the 3' end) in the naturally occurring genome of the organism from which it is derived. The term therefore includes, for example, a recombinant nucleic acid which is incorporated into a vector, into an autonomously replicating plasmid or virus, or into the genomic DNA of a prokaryote or eukaryote; or which exists as a separate molecule (*e.g.*, a cDNA or a genomic DNA fragment produced by PCR or restriction endonuclease treatment) independent of other sequences. It also includes a recombinant DNA which is part of a hybrid gene encoding additional polypeptide sequences and/or including exogenous regulatory elements.

Transformed cell. As used herein, a "transformed cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid molecule of interest. The nucleic acid of interest will typically encode a peptide or protein. The transformed cell may express the sequence of interest or may be used only to propagate the sequence. The term "transformed" may be used herein to embrace any method of introducing exogenous nucleic acids including, but not limited to, transformation, transfection, electroporation, microinjection, viral-mediated transfection, and the like.

Operably joined. As used herein, a coding sequence and a regulatory region are said to be "operably joined" when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory region. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of promoter function results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the regulatory region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a regulatory region would be operably joined to a coding sequence if the regulatory region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

Stringent hybridization conditions. Stringent hybridization conditions is a term of art understood by those of ordinary skill in the art. For any given nucleic acid sequence, stringent hybridization conditions are those conditions of temperature, chaotropic acids, buffer, and ionic

strength which will permit hybridization of that nucleic acid sequence to its complementary sequence and not to substantially different sequences. The exact conditions which constitute "stringent" conditions, depend upon the nature of the nucleic acid sequence, the length of the sequence, and the frequency of occurrence of subsets of that sequence within other non-identical sequences. By varying hybridization conditions from a level of stringency at which non-specific hybridization occurs to a level at which only specific hybridization is observed, one of ordinary skill in the art can, without undue experimentation, determine conditions which will allow a given sequence to hybridize only with complementary sequences. Suitable ranges of such stringency conditions are described in KRAUSE ET AL., METHODS IN ENZYMOLOGY, 200: 546-56 (1991). Stringent hybridization conditions, depending upon the length and commonality of a sequence, may include temperatures of 20°C-65°C and ionic strengths from 5x to 0.1x SSC. Highly stringent hybridization conditions may include temperatures as low as 40-42°C (when denaturants such as formamide are included) or up to 60-65°C in ionic strengths as low as 0.1x SSC. These ranges, however, are only illustrative and, depending upon the nature of the target sequence, and possible future technological developments, may be more stringent than necessary. Less than stringent conditions are employed to isolate nucleic acid sequences which are substantially similar, allelic or homologous to any given sequence.

Selectively bind. As used herein with respect to antibodies, an antibody is said to "selectively bind" to a target if the antibody recognizes and binds the target of interest but does not substantially recognize and bind other molecules in a sample, e.g., a biological sample, which includes the target of interest.

CalDAG-GEF- or cAMP-GEF-associated disorder, condition, or disease. AS used herein, the term "CalDAG-GEF or cAMP-GEF associated disorder, condition, or disease" means any disorder, condition, or disease to which a normal or mutant CalDAG-GEF and/or cAMP-GEF is related in any manner, such as in the causation, prevention, exacerbation, alleviation of the disorder. Thus, as used herein, a CalDAG-GEF- or cAMP-GEF-associated disorder, condition, or disease includes disorders related to the Ras-pathway, such as Ras-related cancers.

Adapter protein. As used herein, the term "adapter protein" means any protein that binds or is bound to a CalDAG-GEF or a cAMP-GEF protein, and facilitates localization of the bound CalDAG-GEF or cAMP-GEF at the plasma membrane, thereby facilitating Ras activation.

Variant. As used herein a “variant” sequence has, or will result in having, a sufficient amino acid similarity to have a reasonable expectation of success in the methods of the present invention. In order to produce variants of the disclosed sequences that may also functionally serve as a CalDAG-GEF or cAMP-GEF protein, any one or more of the naturally-occurring

5 CalDAG-GEF or cAMP-GEF sequences disclosed herein may be used as a reference sequence to determine whether a candidate sequence possesses sufficient amino acid similarity to have a reasonable expectation of success in the methods of the present invention. Preferably, variant sequences are at least 70% similar or 60% identical, more preferably at least 75% similar or 65% identical, and most preferably 80% similar or 70% identical to one of the disclosed, naturally-
10 occurring sequences.

To determine whether a candidate peptide region has the requisite percentage similarity or identity to a reference polypeptide or peptide oligomer, the candidate amino acid sequence and the reference amino acid sequence are first aligned using the dynamic programming algorithm described in Smith and Waterman (1981), J. Mol. Biol. 147:195-197, in combination with the
15 BLOSUM62 substitution matrix described in Figure 2 of Henikoff and Henikoff (1992), “Amino acid substitution matrices from protein blocks”, PNAS (1992 Nov), 89:10915-10919. For the present invention, an appropriate value for the gap insertion penalty is -12, and an appropriate value for the gap extension penalty is -4. Computer programs performing alignments using the algorithm of Smith-Waterman and the BLOSUM62 matrix, such as the GCG program suite
20 (Oxford Molecular Group, Oxford, England), are commercially available and widely used by those skilled in the art.

Once the alignment between the candidate and reference sequence is made, a percent similarity score may be calculated. The individual amino acids of each sequence are compared sequentially according to their similarity to each other. If the value in the BLOSUM62 matrix
25 corresponding to the two aligned amino acids is zero or a negative number, the pairwise similarity score is zero; otherwise the pairwise similarity score is 1.0. The raw similarity score is the sum of the pairwise similarity scores of the aligned amino acids. The raw score is then normalized by dividing it by the number of amino acids in the smaller of the candidate or reference sequences. The normalized raw score is the percent similarity. Alternatively, to
30 calculate a percent identity, the aligned amino acids of each sequence are again compared

sequentially. If the amino acids are non-identical, the pairwise identity score is zero; otherwise the pairwise identity score is 1.0. The raw identity score is the sum of the identical aligned amino acids. The raw score is then normalized by dividing it by the number of amino acids in the smaller of the candidate or reference sequences. The normalized raw score is the percent identity. Insertions and deletions are ignored for the purposes of calculating percent similarity and identity. Accordingly, gap penalties are not used in this calculation, although they are used in the initial alignment.

In all instances, variants of the naturally-occurring CalDAG-GEF or cAMP-GEF proteins, as described above, must be tested for biological activity as described below.

Specifically, the proteins must exhibit guanine nucleotide exchange factor activity, and, preferably, they have the ability to inhibit Ras signaling of the Ras/Raf-1/MAP kinase pathway.

II. The CalDAG-GEFs

CalDAG-GEFI has a substrate specificity for Rap1A, dual binding domains for calcium (Ca^{2+}) and diacylglycerol (DAG), and enriched expression in brain basal ganglia pathways and their axon-terminal regions. Expression of CalDAG-GEFI activates Rap1A and inhibits Ras-dependent activation of the Erk/MAP kinase cascade in 293T cells. Ca^{2+} ionophore and phorbol ester strongly and additively enhance this Rap1A activation. By contrast, CalDAG-GEFII exhibits a different brain expression pattern and fails to activate Rap1A, but activates H-Ras, R-Ras and the Erk/MAP kinase cascade under Ca^{2+} and DAG modulation. The CalDAG-GEF proteins have a critical neuronal function in determining the relative activation of Ras and Rap1 signaling induced by Ca^{2+} and DAG mobilization. The expression of CalDAG-GEFI and CalDAG-GEFII in hematopoietic organs indicates that such control has broad significance in Ras/Rap regulation of normal and malignant states.

The basal ganglia are centrally implicated in movement control and in forms of procedural learning related to habit formation. It is not yet known whether particular neurochemical specializations of the basal ganglia contribute to these behavioral functions. The basal ganglia do, however, have a unique double-inhibitory pathway design combined with abundant expression of neuromodulators in striatal neurons. A number of genes with differentially high expression in the striatum have also been identified. These include genes

coding for proteins with signaling functions, such as adenylate cyclase V (Glatt et al., 361 NATURE (LONDON), 536-38 (1993)) and DARPP-32 (Hemmings et al., 310 NATURE (LONDON) 502-05 (1984)). To identify other cellular signaling molecules that could contribute to basal ganglia functions, a search for striatum-enriched transcripts was performed by a differential

5 display method, as discussed in Example 1. Among the transcripts identified in this search were a family of genes characterized by the presence of a Ras superfamily (GEF) domain.

Specific domains identified include structurally conserved GEF regions SCR1, SCR2, and SCR3, as shown in Figures 2C and 3D, and as shown in the following table.

TABLE 1

Gene	SCR1	SCR2	SCR3
hCalDAG-GEFI	SEQ ID NO.3: 605-677 SEQ ID NO.4: 149-173	SEQ ID NO.3: 817-946 SEQ ID NO.4: 219-262	SEQ ID NO.3: 1053-1185 SEQ ID NO.4: 298-320
hCalDAG-GEFII	SEQ ID NO.7: 728-800 SEQ ID NO.8: 205-229	SEQ ID NO.7: 913-1042 SEQ ID NO.8: 270-313	SEQ ID NO.7: 1084-1216 SEQ ID NO.8: 348-371
hcAMP-GEFI	SEQ ID NO.11: 2058-2130 SEQ ID NO.12: 205-229	SEQ ID NO.11: 2276-2405 SEQ ID NO.12: 688-731	SEQ ID NO.11: 2516-2582 SEQ ID NO.12: 767-789
rcAMP-GEFI	SEQ ID NO.9: 2050-2122 SEQ ID NO.10: 618-642	SEQ ID NO.9: 2267-2396 SEQ ID NO.10: 691-734	SEQ ID NO.9: 2502-2568 SEQ ID NO.10: 770-792
hcAMP-GEFII	SEQ ID NO.17: 2707-2779 SEQ ID NO.18: 767-791		
rcAMP-GEFII	SEQ ID NO.15: 576-648 SEQ ID NO.16: 192-216		

In addition, the EF hand and DAG-binding domains were identified as shown in Figures 2E and 2F, and as shown in the following table:

Table 2

Gene	EF Hand Domain	DAG-Binding Domain
hCalDaG-GEFI	SEQ ID NO.3: 1456-1516 SEQ ID NO.4: 432-452	SEQ ID NO.3: 1652-1804 SEQ ID NO.4: 498-548
hCalDAG-GEFII	SEQ ID NO.7: 1384-1444 SEQ ID NO.8: 427-447	SEQ ID NO.7: 1579-1729 SEQ ID NO.8: 492-542

5

Finally, the cAMP-binding domains were identified as shown in Figure 3E, and as shown in the following table:

Table 3

Gene	cAMP-Binding Domain
hcAMP-GEFI	SEQ ID NO.11: 2012-2255 SEQ ID NO.12: 219-300
rcAMP-GEFI	SEQ ID NO.9: 853-1096 SEQ ID NO.10: 219-300
rcAMP-GEFII	SEQ ID NO.17: 1522-1765 SEQ ID NO.18: 372-453

III. The cAMP-GEFs

Cyclic adenosine 3', 5'-monophosphate (cAMP) is a universal second messenger that induces a variety of physiological responses in eukaryotic cells ranging from growth, differentiation, and gene expression to secretion and neurotransmission. The cAMP second messenger system has also been centrally implicated in modulating synaptic function, neuroplasticity and learning and memory. Most of these effects have been attributed to the binding of cAMP to cAMP-dependent protein kinase (PKA), leading in turn to the activation of intracellular phosphorylation cascades. Reported herein is the identification of a new family of cAMP binding proteins that are differentially distributed in the brain and body organs and that are characterized by the presence of both a cAMP binding domain and a guanine nucleotide exchange factor (GEF) domain. These proteins, cAMP-GEFs, bind cAMP and selectively activate the Ras superfamily small G protein, Rap1A, in a cAMP-dependent but PKA-independent manner.

The general concept of cAMP signaling involves the sequential activation (or inhibition) of cAMP production by G proteins, the binding of cAMP to PKA, and the triggering of a series of downstream serine-threonine phosphorylation cascades. Viewed as the nearly exclusive target of cAMP binding in eukaryotic cells, PKA has been considered the essential effector molecule mediating a wide range of physiological effects of G protein/cAMP-triggered phosphorylation cascades. As the main cAMP effector, PKA has also been shown to function in the indirect coupling of the cAMP signal transduction system to other intracellular signaling cascades. The cAMP signaling system has also been strongly implicated in neuronal functions ranging from neurotransmitter-initiated signaling to neuroplasticity underlying development and memory, but PKA has not been clearly linked to all of these neuronal functions, and region-specific neuronal effects have been observed as well. The cAMP-GEF gene has a Ras superfamily GEF motif. Thus, the gene codes for a novel cAMP binding protein that directly couples the cAMP signal transduction system to Ras superfamily cascades.

IV. Preferred Embodiments

Based, in part, upon the discoveries disclosed and described herein, the following preferred embodiments of the present invention are provided.

1. Isolated Nucleic Acids

In one series of embodiments, the present invention provides isolated nucleic acids corresponding to, or relating to, the CalDAG-GEF or cAMP-GEF nucleic acid sequences disclosed herein. As described more fully below, these sequences include normal CalDAG-GEF and cAMP-GEF sequences from humans and other mammalian species, mutant CalDAG-GEF and cAMP-GEF sequences from humans and other mammalian species, homologous sequences from non-mammalian species such as *Drosophila* and *C. elegans*, subsets of these sequences useful as probes and PCR primers, subsets of these sequences encoding fragments of the CalDAG-GEF or cAMP-GEF proteins or corresponding to particular domains or polymorphic regions, complementary or antisense sequences corresponding to fragments of the CalDAG-GEF or cAMP-GEF genes, sequences in which the CalDAG-GEF and/or cAMP-GEF coding regions have been operably joined to exogenous regulatory regions, and sequences encoding fusion proteins of the portions of the CalDAG-GEF or cAMP-GEF proteins fused to other proteins useful as markers of expression, as "tags" for purification, or in screens and assays for proteins interacting with the CalDAG-GEFs and/or cAMP-GEFs.

Thus, in a first series of embodiments, isolated nucleic acid sequences are provided which encode normal versions of the CalDAG-GEF and cAMP-GEF proteins. Examples of such nucleic acid sequences are disclosed herein. These nucleic acids may be genomic sequences or may be cDNA sequences (*e.g.*, SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, and 17). Thus, for example, the invention provides nucleic acid sequences in which the alternative splice variants described herein are incorporated at the DNA level, thereby, enabling cells including these sequences to express only one of the alternative splice variants at each splice position. For example, a recombinant gene may be produced in which one of the splice variants of cAMP-GEF is incorporated into DNA such that cells having this recombinant gene can express only one of these variants. For purposes of reducing the size of a recombinant CalDAG-GEF or cAMP-GEF gene, a cDNA gene may be employed or various combinations of the introns and untranslated exons may be removed from a DNA construct. Such constructs may be particularly useful, as described below, in identifying compounds which can induce or repress the expression of the CalDAG-GEF or cAMP-GEF genes.

In addition to the disclosed CalDAG-GEF and/or cAMP-GEF sequences, one of ordinary skill in the art is now enabled to identify and isolate nucleic acids corresponding to CalDAG-GEF or cAMP-GEF genes or cDNAs which are allelic to the disclosed sequences or which are heterospecific homologues. Thus, the present invention provides isolated nucleic acids

5 corresponding to these alleles and homologues, as well as various recombinant constructs derived from these sequences, by means which are well known in the art. Briefly, one of ordinary skill in the art may now screen preparations of genomic or cDNA, including samples prepared from individual organisms (*e.g.*, human cancer patients or their family members) as well as bacterial, viral, yeast or other libraries of genomic or cDNA, using probes or PCR

10 primers to identify allelic or homologous sequences. Because it is desirable to identify additional CalDAG-GEF and/or cAMP-GEF gene mutations which may contribute to the development of Ras-related cancers, because it is desirable to identify additional CalDAG-GEF and/or cAMP-GEF polymorphisms which are not mutant or have antitumorigenic effects, and because it is also desired to create a variety of animal models which may be used to study Ras-related cancers and

15 screen for potential therapeutics, it is particularly contemplated that additional CalDAG-GEF and/or cAMP-GEF sequences will be isolated from other preparations or libraries of human nucleic acids and from preparations or libraries from animals including rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates. Furthermore, CalDAG-GEF and/or cAMP-GEF homologues from yeast or invertebrate species, including *C.*

20 *elegans* and other nematodes, as well as *Drosophila* and other insects, may have particular utility for drug screening. For example, invertebrates bearing mutant CalDAG-GEF and/or cAMP-GEF homologues (or mammalian CalDAG-GEF and/or cAMP-GEF transgenes) which cause a rapidly occurring and easily scored phenotype (*e.g.*, abnormal eye development after several days) can be used as screens for drugs which block the effect of the mutant gene. Such invertebrates may

25 prove far more rapid and efficient for mass screenings than larger vertebrate animals. Once lead compounds are found through such screens, they may be tested in higher animals.

Depending upon the intended use, the present invention provides nucleic acid subsequences of the CalDAG-GEF and/or cAMP-GEF genes which may have lengths varying from 8-10 nucleotides (*e.g.*, for use as PCR primers) to nearly the full size of the CalDAG-GEF

30 and/or cAMP-GEF genes. Thus, the present invention provides isolated nucleic acids

comprising sequences corresponding to at least 8 , preferably at least 10, and more preferably at least 15 consecutive, nucleotides of the CalDAG-GEF and/or cAMP-GEF genes, as disclosed or otherwise enabled herein, or to their complements.

In another series of embodiments, the present invention provides for isolated nucleic acids encoding all or a portion of the CalDAG-GEF and/or cAMP-GEF proteins in the form of a fusion protein. In these embodiments, a nucleic acid regulatory region (endogenous or exogenous) is operably joined to a first coding region which is covalently joined in-frame to a second coding region. The CalDAG-GEF and/or cAMP-GEF sequences of the fusion protein may represent the first, second, or any additional coding regions. The CalDAG-GEF and/or cAMP-GEF sequences may be conserved or non-conserved domains and can be placed in any coding region for the fusion protein.

In another series of embodiments, the present invention provides isolated nucleic acids in the form of recombinant DNA constructs in which a marker or reporter gene (*e.g.*, β -galactosidase, luciferase) is operably joined to the 5' regulatory region of a CalDAG-GEF and/or cAMP-GEF gene such that expression of the marker gene is under the control of the CalDAG-GEF and/or cAMP-GEF regulatory sequences. Such isolated nucleic acids may be used to produce cells, cell lines or transgenic animals which are useful in the identification of compounds which can, directly or indirectly, differentially affect the expression of the CalDAG-GEFs and/or cAMP-GEFs.

Finally, the isolated nucleic acids of the present invention include any of the above described sequences when included in vectors. Appropriate vectors include cloning vectors and expression vectors of all types, including plasmids, phagemids, cosmids, episomes, and the like, as well as integration vectors. The vectors may also include various marker genes (*e.g.*, antibiotic resistance or susceptibility genes) which are useful in identifying cells successfully transformed therewith. In addition, the vectors may include regulatory sequences to which the nucleic acids of the invention are operably joined, and/or may also include coding regions such that the nucleic acids of the invention, when appropriately ligated into the vector, are expressed as fusion proteins. Such vectors may also include vectors for use in yeast "two hybrid," baculovirus, and phage-display systems.

2. Substantially Pure Proteins

The present invention provides for substantially pure preparations of the CalDAG-GEF and/or cAMP-GEF proteins, fragments of the CalDAG-GEF and/or cAMP-GEF proteins, and fusion proteins including the CalDAG-GEFs and/or cAMP-GEFs or fragments thereof. The proteins, fragments and fusions have utility, as described herein, in the generation of antibodies to normal and mutant CalDAG-GEFs and/or cAMP-GEFs, in the identification of CalDAG-GEF and/or cAMP-GEF binding proteins, and in diagnostic and therapeutic methods. Therefore, depending upon the intended use, the present invention provides substantially pure proteins or peptides comprising amino acid sequences which are subsequences of the complete CalDAG-GEF and/or cAMP-GEF proteins and which may have lengths varying from 4-8 amino acids (*e.g.*, for use as immunogens), or 9-15 amino acids (*e.g.*, for use in binding assays), to the complete CalDAG-GEF and/or cAMP-GEF proteins. Thus, the present invention provides substantially pure proteins or peptides comprising sequences corresponding to at least 4, preferably at least 9, more preferably at least 15 consecutive amino acids of the CalDAG-GEF and/or cAMP-GEF proteins, as disclosed or otherwise enabled herein.

Purification can be achieved using standard protein purification procedures including, but not limited to, gel-filtration chromatography, ion-exchange chromatography, high-performance liquid chromatography (RP-HPLC, ion-exchange HPLC, size-exclusion HPLC, high-performance chromatofocusing chromatography, hydrophobic interaction chromatography, immunoprecipitation, or immunoaffinity purification. Gel electrophoresis (*e.g.*, PAGE, SDS-PAGE) can also be used to isolate a protein or peptide based on its molecular weight, charge properties, and hydrophobicity.

A CalDAG-GEF or cAMP-GEF protein, or a fragment thereof, may also be conveniently purified by creating a fusion protein including the desired CalDAG-GEF or cAMP-GEF sequence fused to another peptide such as an antigenic determinant or poly-His tag (*e.g.*, QIAexpress vectors, (QIAGEN Corp., Chatsworth, CA)), or a larger protein (*e.g.*, GST using the pGEX-27 vector (Amrad, USA) or green fluorescent protein using the Green Lantern vector (GIBCO/BRL, Gaithersburg, MD)).

3. Antibodies to the CalDAG-GEF and/or cAMP-GEFs

The present invention also provides antibodies, and methods of making antibodies, which selectively bind to the CalDAG-GEF and/or cAMP-GEF proteins or fragments thereof. The antibodies of the invention have utility as laboratory reagents for, *inter alia*, immunoaffinity
5 purification of the CalDAG-GEFs and/or cAMP-GEFs, Western blotting to identify cells or tissues expressing the CalDAG-GEFs and/or cAMP-GEFs, and immunocytochemistry or immunofluorescence techniques to establish the subcellular location of the protein.

The antibodies of the invention may be generated in a host using the entire CalDAG-GEF and/or cAMP-GEF proteins of the invention or using any CalDAG-GEF and/or cAMP-GEF
10 epitope which is characteristic of that protein and which substantially distinguishes it from host proteins. Such epitopes may be identified by comparing sequences of, for example, 4-8 amino acid residues from a CalDAG-GEF and/or cAMP-GEF sequence to computer databases of protein sequences from the relevant host. Antibodies against highly conserved domains are expected to have the greatest utility for purification or identification of CalDAG-GEFs and/or
15 cAMP-GEFs.

Amino acid residue positions which are potential antigenic sites in the CalDAG-GEF or cAMP-GEF proteins and which may be useful in generating the antibodies of the invention may be determined by using computer programs such as the IBI Pustell program. Other methods of choosing antigenic determinants are known in the art and may, of course, be employed. In
20 addition, larger fragments (*e.g.* 9-15 residues) including some of these epitopes may also be employed. Even larger fragments, including, for example, entire functional domains or multiple functional domains may also be preferred. For an overview of antibody techniques, see Antibody Engineering: A Practical Guide, Borrebaek, ed., W.H. Freeman & Company, NY (1992), or Antibody Engineering, 2nd Ed., Borrebaek, ed., Oxford University Press, Oxford
25 (1995).

The antibodies of the invention may be labelled or conjugated with other compounds or materials for diagnostic and/or therapeutic uses. For example, they may be coupled to radionuclides, fluorescent compounds, or enzymes for imaging or therapy, or to liposomes for the targeting of compounds contained in the liposomes to a specific tissue location.

4. Transformed Cell Lines

The present invention also provides for cells or cell lines, both prokaryotic and eukaryotic, which have been transformed or transfected with the nucleic acids of the present invention so as to cause clonal propagation of those nucleic acids and/or expression of the proteins or peptides encoded thereby. Such cells or cell lines will have utility not only in the propagation and production of the nucleic acids and proteins of the present invention but also, as further described herein, as model systems for diagnostic and therapeutic assays. As used herein, the term "transformed cell" is intended to embrace any cell, or the descendant of any cell, into which has been introduced any of the nucleic acids of the invention, whether by transformation, transfection, infection, electroporation, microinjection or other means. Methods of producing appropriate vectors, transforming cells with those vectors, and identifying transformants are well known in the art and are only briefly reviewed here (see, for example, Sambrook et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

Vectors may be introduced into the recipient or "host" cells by various methods well known in the art including, but not limited to, calcium phosphate transfection, strontium phosphate transfection, DEAE dextran transfection, electroporation, lipofection (*e.g.*, Dosper Liposomal transfection reagent, Boehringer Mannheim, Germany), microinjection, ballistic insertion on micro-beads, protoplast fusion or, for viral or phage vectors, by infection with the recombinant virus or phage.

5. Transgenic Animal Models

The present invention also provides for the production of transgenic non-human animal models for the study of Ras-related cancers, for the screening of candidate pharmaceutical compounds, for the creation of explanted mammalian cell cultures (*e.g.*, neuronal, glial, organotypic or mixed cell cultures) in which mutant or wild type CalDAG-GEF and/or cAMP-GEF sequences are expressed or in which the CalDAG-GEF and/or cAMP-GEF genes have been inactivated (*e.g.*, "knock-out" deletions), and for the evaluation of potential therapeutic interventions.

Species suitable for use as animal models in the present invention include, but are not limited to, rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates (*e.g.*, Rhesus monkeys, chimpanzees).

Various techniques for generating transgenic animals, as well as techniques for homologous recombination or gene targeting, are now widely accepted and practiced. See, for example, Hogan et al., *Manipulating Mouse Embryo* (1986). To create a transgene, the target sequence of interest (*e.g.*, mutant or wild-type CalDAG-GEF or cAMP-GEF sequences) is typically ligated into a cloning site located downstream of a promoter element which will regulate the expression of RNA from the CalDAG-GEF or cAMP-GEF sequence. An alternate approach to creating a transgene is to use endogenous CalDAG-GEF or cAMP-GEF regulatory sequences to drive expression of the CalDAG-GEF or cAMP-GEF transgene.

6. Assays for Drugs Which Affect CalDAG-GEF and/or cAMP-GEF Expression

In another series of embodiments, the present invention provides assays for identifying small molecules or other compounds which are capable of inducing or inhibiting the expression of the CalDAG-GEF or cAMP-GEF genes and proteins. The assays may be performed *in vitro* using non-transformed cells, immortalized cell lines, or recombinant cell lines, or *in vivo* using the transgenic animal models enabled herein.

In particular, the assays may detect the presence of increased or decreased expression of CalDAG-GEF, cAMP-GEF, or other CalDAG-GEF or cAMP-GEF-related genes or proteins, on the basis of increased or decreased mRNA expression (using, *e.g.*, the nucleic acid probes disclosed and enabled herein), increased or decreased levels of CalDAG-GEF, cAMP-GEF or other CalDAG-GEF or cAMP-GEF-related protein products (using, *e.g.*, the anti-CalDAG-GEF or anti-cAMP-GEF antibodies disclosed and enabled herein), or increased or decreased levels of expression of a marker gene (*e.g.*, β -galactosidase or luciferase) operably joined to a CalDAG-GEF or cAMP-GEF 5' regulatory region in a recombinant construct.

Thus, for example, one may culture cells known to express a particular CalDAG-GEF or cAMP-GEF and add to the culture medium one or more test compounds. After allowing a sufficient period of time (*e.g.*, 0-72 hours) for the compound to induce or inhibit the expression of the CalDAG-GEF or cAMP-GEF, any change in levels of expression from an established

baseline may be detected using any of the techniques described above and well known in the art. In particularly preferred embodiments, the cells are from an immortalized cell line such as a human neuroblastoma, glioblastoma or a hybridoma cell line. Using the nucleic acid probes and/or antibodies disclosed and enabled herein, detection of changes in the expression of a

5 CalDAG-GEF or cAMP-GEF and thus, identification of the compound as an inducer or repressor of CalDAG-GEF and/or cAMP-GEF expression, requires only routine experimentation.

In particularly preferred embodiments, a recombinant assay is employed in which a reporter gene such as a β -galactosidase, green fluorescent protein, alkaline phosphatase, or luciferase is operably joined to a 5' regulatory region of a CalDAG-GEF or cAMP-GEF gene.

10 The reporter gene and regulatory regions are joined in-frame (or in each of the three possible reading frames) so that transcription and translation of the reporter gene may proceed under the control of the CalDAG-GEF or cAMP-GEF regulatory elements. The recombinant construct may then be introduced into any appropriate cell type, although mammalian cells are preferred, and human cells are most preferred. The transformed cells may be grown in culture and, after
15 establishing the baseline level of expression of the reporter gene, test compounds may be added to the medium. The ease of detection of the expression of the reporter gene provides for a rapid, high through-put assay for the identification of inducers and repressors of the CalDAG-GEF or cAMP-GEF gene.

Compounds identified by this method will have potential utility in modifying the
20 expression of the CalDAG-GEF, cAMP-GEF or other CalDAG-GEF or cAMP-GEF-related genes *in vivo*. These compounds may be further tested in the animal models disclosed and enabled herein to identify those compounds having the most potent *in vivo* effects. In addition, as described herein with respect to small molecules having CalDAG-GEF or cAMP-GEF-binding activity, these molecules may serve as "lead compounds" for the further development of
25 pharmaceuticals by, for example, subjecting the compounds to sequential modifications, molecular modeling, and other routine procedures employed in rational drug design.

7. Identification of Compounds with CalDAG-GEF and/or cAMP-GEF Binding Capacity

In light of the present disclosure, one of ordinary skill in the art is enabled to practice new
30 screening methodologies which will be useful in the identification of proteins and other

compounds which bind to, or otherwise directly interact with, the CalDAG-GEFs or cAMP-GEFs. The proteins and compounds will include endogenous cellular components which interact with the CalDAG-GEFs or cAMP-GEFs *in vivo* and which, therefore, provide new targets for pharmaceutical and therapeutic interventions, as well as recombinant, synthetic and otherwise exogenous compounds which may have CalDAG-GEF or cAMP-GEF binding capacity and, therefore, may be candidates for pharmaceutical agents. Thus, in one series of embodiments, cell lysates or tissue homogenates (*e.g.*, human brain homogenates, leukocyte lysates) may be screened for proteins or other compounds which bind to one of the normal or mutant CalDAG-GEFs and/or cAMP-GEFs. Alternatively, any of a variety of exogenous compounds, both naturally occurring and/or synthetic (*e.g.*, libraries of small molecules or peptides), may be screened for CalDAG-GEF or cAMP-GEF binding capacity. Small molecules are particularly preferred in this context because they are more readily absorbed after oral administration, have fewer potential antigenic determinants, and/or are more likely to cross the blood brain barrier than larger molecules such as nucleic acids or proteins. The methods of the present invention are particularly useful in that they may be used to identify molecules which selectively or preferentially bind to a mutant form of a CalDAG-GEF or cAMP-GEF protein (rather than a normal form) and, therefore, may have particular utility in treating the heterozygous victims of a CalDAG-GEF or cAMP-GEF associated disorder.

Compounds which bind to normal, mutant or both forms of the CalDAG-GEFs or cAMP-GEFs may have utility in treatments and diagnostics. Compounds which bind only to a normal CalDAG-GEF or cAMP-GEF may, for example, act as enhancers of its normal activity and thereby at least partially compensate for the lost or abnormal activity of mutant forms of the CalDAG-GEF or cAMP-GEF in victims suffering from CalDAG-GEF- or cAMP-GEF-associated disorders. Compounds which bind to both normal and mutant forms of a CalDAG-GEF or cAMP-GEF may have utility if they differentially affect the activities of the two forms so as to alleviate the overall departure from normal function. Alternatively, blocking the activity of both normal and mutant forms of either CalDAG-GEF or cAMP-GEF may have less severe physiological and clinical consequences than the normal progress of the disorder and, therefore, compounds which bind to and inhibit the activity of both normal and mutant forms of a CalDAG-GEF or cAMP-GEF may be therapeutically useful. Preferably, however, compounds

are identified which have a higher affinity of binding to mutant CalDAG-GEF or cAMP-GEF than to normal CalDAG-GEF or cAMP-GEF, and which selectively or preferentially inhibit the activity of the mutant form. Such compounds may be identified by using any of the techniques described herein, and then comparing the binding affinities of the candidate compound(s) for the normal and mutant forms of CalDAG-GEF or cAMP-GEF.

The effect of agents which bind to the CalDAG-GEFs or cAMP-GEFs (normal or mutant forms of either) can be monitored either by direct monitoring of this binding (*e.g.*, using the BIAcore assay, LKB Pharmacia, Sweden) or by indirect monitoring of binding by detecting, for example, a change in fluorescence, molecular weight, or concentration of either the binding agent or CalDAG-GEF or cAMP-GEF component, either in a soluble phase or in a substrate-bound phase.

Once identified by the methods described above, the candidate compounds may then be produced in quantities sufficient for pharmaceutical administration or testing (*e.g.*, μg or mg or greater quantities), and formulated in a pharmaceutically acceptable carrier (see, *e.g.*, REMINGTON'S PHARMACEUTICAL SCIENCES, Gennaro, A., ed., Mack Pub., (1990)). These candidate compounds may then be administered to the transformed cells of the invention, to the transgenic animal models of the invention, to cell lines derived from the animal models or from human patients, or to patients with CalDAG-GEF- or cAMP-GEF-associated disorders. The animal models described and enabled herein are of particular utility in further testing candidate compounds which bind to normal or mutant CalDAG-GEF or cAMP-GEF for their therapeutic efficacy.

In addition, once identified by the methods described above, the candidate compounds may also serve as "lead compounds" in the design and development of new pharmaceuticals. For example, as is well known in the art, sequential modification of small molecules (*e.g.*, amino acid residue replacement for peptides; functional group replacement for peptide or non-peptide compounds) is a standard approach in the pharmaceutical industry for the development of new pharmaceuticals. Such development generally proceeds from a "lead compound" which is shown to have at least some of the activity (*e.g.*, CalDAG-GEF or cAMP-GEF binding or blocking ability) of the desired pharmaceutical. In particular, when one or more compounds having at least some activity of interest (*e.g.*, modulation of CalDAG-GEF or cAMP-GEF activity) are

identified, structural comparison of the molecules can greatly inform the skilled practitioner by suggesting portions of the lead compounds which should be conserved and portions which may be varied in the design of new candidate compounds. Thus, the present invention also provides a means of identifying lead compounds which may be sequentially modified to produce new candidate compounds for use in the treatment or CalDAG-GEF- or cAMP-GEF-associated disorders. These new compounds then may be tested both for CalDAG-GEF or cAMP-GEF-binding or blocking (*e.g.*, in the binding assays described above) and for therapeutic efficacy (*e.g.*, in the animal models described herein). This procedure may be iterated until compounds having the desired therapeutic activity and/or efficacy are identified.

10 In each of the present series of embodiments, an assay is conducted to detect binding between a "CalDAG-GEF component" or a "cAMP-GEF component" and some other moiety. Of particular utility will be sequential assays in which compounds are tested for the ability to bind to only the normal or only the mutant forms of the CalDAG-GEF or cAMP-GEF functional domains using mutant and normal CalDAG-GEF or cAMP-GEF components in the binding
15 assays. Such compounds are expected to have the greatest therapeutic utilities, as described more fully below. The "CalDAG-GEF component" or the "cAMP-GEF component" in these assays may be a complete normal or mutant form of a CalDAG-GEF or cAMP-GEF protein (*e.g.*, an hCalDAG-GEF or hcAMP-GEF variant) but need not be, or a specific domain of a CalDAG-GEF or cAMP-GEF.. Rather, particular functional domains of the CalDAG-GEFs or cAMP-
20 GEFs, as described above, may be employed either as separate molecules or as part of a fusion protein. For example, to isolate proteins or compounds that interact with these functional domains, screening may be carried out using fusion constructs and/or synthetic peptides corresponding to these regions. Obviously, various combinations of fusion proteins and functional domains from CalDAG-GEF or cAMP-GEF are possible. In addition, the functional
25 domains may be altered so as to aid in the assay by, for example, introducing into the functional domain a reactive group or amino acid residue (*e.g.*, cysteine) which will facilitate immobilization of the domain on a substrate (*e.g.*, using sulfhydryl reactions).

Methods for screening cellular lysates, tissue homogenates, or small molecule libraries for candidate CalDAG-GEF or cAMP-GEF-binding molecules are well known in the art and, in
30 light of the present disclosure, may now be employed to identify compounds which bind to

normal or mutant CalDAG-GEF or cAMP-GEF components or which modulate CalDAG-GEF or cAMP-GEF activity as defined by non-specific measures (*e.g.*, changes in intracellular Ca^{2+} , GTP/GDP ratio) or by specific measures (*e.g.*, changes in the expression of other downstream genes which can be monitored by differential display, 2D gel electrophoresis, differential hybridization, or SAGE methods). The preferred methods involve variations on the following techniques: (1) direct extraction by affinity chromatography; (2) co-isolation of CalDAG-GEF or cAMP-GEF components and bound proteins or other compounds by immunoprecipitation; (3) the Biomolecular Interaction Assay (BIAcore); and (4) the yeast two-hybrid systems. These and others are discussed separately below.

A. Affinity Chromatography

In light of the present disclosure, a variety of affinity binding techniques well known in the art may be employed to isolate proteins or other compounds which bind to the CalDAG-GEFs or cAMP-GEFs disclosed or otherwise enabled herein. In general, a CalDAG-GEF or cAMP-GEF component may be immobilized on a substrate (*e.g.*, a column or filter) and a solution including the test compound(s) is contacted with the CalDAG-GEF or cAMP-GEF protein, fusion or fragment under conditions which are permissive for binding. The substrate is then washed with a solution to remove unbound or weakly bound molecules. A second wash may then elute those compounds which strongly bound to the immobilized normal or mutant CalDAG-GEF or cAMP-GEF component. Alternatively, the test compounds may be immobilized and a solution containing one or more CalDAG-GEF or cAMP-GEF components may be contacted with the column, filter, or other substrate. The ability of the CalDAG-GEF or cAMP-GEF component to bind to the test compounds may be determined as above or a labeled form of the CalDAG-GEF or cAMP-GEF component (*e.g.*, a radio-labeled or chemiluminescent functional domain) may be used to more rapidly assess binding to the substrate-immobilized compound(s).

B. Co-Immunoprecipitation

Another well characterized technique for the isolation of the CalDAG-GEF or cAMP-GEF components and their associated proteins or other compounds is direct immunoprecipitation with antibodies. This procedure has been successfully used, for example, to isolate many of the synaptic vesicle associated proteins (Phizicky et al., 59 J. BIOL. CHEM. 94-123 (1994)). Thus,

either normal or mutant CalDAG-GEF or cAMP-GEF components may be mixed in a solution with the candidate compound(s) under conditions which are permissive for binding, and the CalDAG-GEF or cAMP-GEF component may be immunoprecipitated. Proteins or other compounds which co-immunoprecipitate with the CalDAG-GEF or cAMP-GEF component may then be identified by standard techniques as described above. General techniques for immunoprecipitation may be found in, for example, Harlow et al., ANTIBODIES: A LABORATORY MANUAL (1988).

The antibodies employed in this assay, as described and enabled herein, may be polyclonal or monoclonal, and include the various antibody fragments as well as single chain antibodies, and the like.

C. The Biomolecular Interaction Assay

Another useful method for the detection and isolation of binding proteins is the Biomolecular Interaction Assay or "BIAcore" system developed by Pharmacia Biosensor and described in the manufacturer's protocol (LKB Pharmacia, Sweden). In light of the present disclosure, one of ordinary skill in the art is now enabled to employ this system, or a substantial equivalent, to identify proteins or other compounds having CalDAG-GEF or cAMP-GEF binding capacity. The BIAcore system uses an affinity purified anti-GST antibody to immobilize GST-fusion proteins onto a sensor chip. Obviously, other fusion proteins and corresponding antibodies may be substituted. The sensor utilizes surface plasmon resonance which is an optical phenomenon that detects changes in refractive indices. A homogenate of a tissue of interest is passed over the immobilized fusion protein and protein-protein interactions are registered as changes in the refractive index. This system can be used to determine the kinetics of binding and to assess whether any observed binding is of physiological relevance.

D. The Yeast Two-Hybrid System

The yeast "two-hybrid" system takes advantage of transcriptional factors that are composed of two physically separable, functional domains. One commonly used system employs the yeast GAL4 transcriptional activator, consisting of a DNA binding domain and a transcriptional activation domain. Two different cloning vectors are used to generate separate fusions of the GAL4 domains to genes encoding potential binding proteins. The fusion proteins

are co-expressed, targeted to the nucleus and, if interactions occur, activation of a reporter gene (*e.g.*, lacZ) produces a detectable phenotype.

E. Other Methods

The nucleotide sequences and protein products, including both mutant and normal forms of these nucleic acids and their corresponding proteins, can be used with the above techniques to isolate other interacting proteins, and to identify other genes whose expression is altered by the over-expression of normal CalDAG-GEF or cAMP-GEF sequences, by the under-expression of normal CalDAG-GEFs or cAMP-GEFs sequences, or by the expression of mutant CalDAG-GEF and/or cAMP-GEF sequences. Identification of these interacting proteins, as well as the identification of other genes whose expression levels are altered in the face of mutant CalDAG-GEF or cAMP-GEF sequences (for instance) will identify other gene targets which have direct relevance to the pathogenesis of this disease in its clinical or pathological forms. Specifically, these techniques rely on PCR-based and/or hybridization-based methods to identify genes which are differentially expressed between two conditions (a cell line expressing normal CalDAG-GEFs or cAMP-GEFs compared to the same cell type expressing a mutant CalDAG-GEF or cAMP-GEF sequence). These techniques include differential display, serial analysis of gene expression (SAGE), mass-spectrometry of protein, 2D-gels and subtractive hybridization (*See, e.g.*, Nowak, 270 Sci. 368-371 (1995); Kahn, 270 Sci. 369-370 (1995)).

8. Methods of Identifying Compounds Modulating CalDAG-GEF and/or cAMP-GEF Activity

In another series of embodiments, the present invention provides for methods of identifying compounds with the ability to modulate the activity of normal and mutant CalDAG-GEFs and/or cAMP-GEFs. As used with respect to this series of embodiments, the term “activity” broadly includes gene and protein expression, CalDAG-GEF and/or cAMP-GEF protein post-translation processing, trafficking and localization, and any functional activity (*e.g.*, enzymatic, receptor-effector, binding, channel), as well as downstream affects of any of these. Using the transformed cells and transgenic animal models of the present invention, cells obtained from subjects bearing a mutant CalDAG-GEF and/or cAMP-GEF gene, or animals or human subjects bearing naturally occurring CalDAG-GEF and/or cAMP-GEF mutations, it is now possible to screen candidate pharmaceuticals and treatments for their therapeutic effects by

detecting changes in one or more of the functional characteristics or phenotypic manifestations of normal or mutant CalDAG-GEF and/or cAMP-GEF expression.

Thus, the present invention provides methods for screening or assaying for proteins, small molecules or other compounds which modulate CalDAG-GEF and/or cAMP-GEF activity by contacting a cell *in vivo* or *in vitro* with a candidate compound and assaying for a change in a marker associated with normal or mutant CalDAG-GEF and/or cAMP-GEF activity. The marker associated with CalDAG-GEF and/or cAMP-GEF activity may be any measurable biochemical, physiological, histological and/or behavioral characteristic associated with CalDAG-GEF and/or cAMP-GEF expression. In particular, useful markers will include any measurable biochemical, physiological, histological and/or behavioral characteristic which distinguishes cells, tissues, animals or individuals bearing at least one mutant CalDAG-GEF and/or cAMP-GEF gene from their normal counterparts. In addition, the marker may be any specific or non-specific measure of CalDAG-GEF and/or cAMP-GEF activity, such as the GDP/GTP bound to Rap1/Ras. CalDAG-GEF and/or cAMP-GEF specific measures include measures of CalDAG-GEF and/or cAMP-GEF expression (*e.g.*, CalDAG-GEF and/or cAMP-GEF mRNA or protein levels) which may employ the nucleic acid probes or antibodies of the present invention. Non-specific measures include changes in cell physiology such as pH, intracellular calcium, cAMP levels, overall GTP/GDP ratios, phosphatidylinositol activity, protein phosphorylation, etc., which can be monitored by known methods. The activation or inhibition of CalDAG-GEF or cAMP-GEF activity in its mutant or normal form can also be monitored by examining changes in the expression of other genes which are specific to the CalDAG-GEF and/or cAMP-GEF pathway. These can be assayed by such techniques as differential display, differential hybridization, and SAGE, as well as by 2-D gel electrophoresis of cellular lysates. In each case, the differentially-expressed genes can be ascertained by inspection of identical studies before and after application of the candidate compound. Furthermore, as noted elsewhere, the particular genes whose expression is modulated by the administration of the candidate compound can be ascertained by cloning, nucleotide sequencing, amino acid sequencing, or mass spectrometry.

In general, a cell may be contacted with a candidate compound and, after an appropriate period (*e.g.*, 0-72 hours for most biochemical measures of cultured cells), the marker of CalDAG-GEF or cAMP-GEF activity may be assayed and compared to a baseline measurement.

The baseline measurement may be made prior to contacting the cell with the candidate compound or may be an external baseline established by other experiments or known in the art. The cell may be a transformed cell of the present invention or an explant from an animal or individual. In particular, the cell may be an explant from a carrier of a CalDAG-GEF or cAMP-GEF mutation or an animal model of the invention (*e.g.*, a transgenic nematode or mouse bearing a mutant CalDAG-GEF or cAMP-GEF gene). Preferred cells include those from neurological tissues such as neuronal, glial or mixed cell cultures; and cultured fibroblasts, liver, kidney, spleen, or bone marrow. The cells may be contacted with the candidate compounds in a culture *in vitro* or may be administered *in vivo* to a live animal or human subject. For live animals or human subjects, the test compound may be administered orally or by any parenteral route suitable to the compound. For clinical trials of human subjects, measurements may be conducted periodically (*e.g.*, daily, weekly or monthly) for several months or years.

In light of the identification, characterization, and disclosure herein of the CalDAG-GEF or cAMP-GEF genes and proteins, the CalDAG-GEF or cAMP-GEF nucleic acid probes and antibodies, and the CalDAG-GEF or cAMP-GEF transformed cells and transgenic animals of the invention, one of ordinary skill in the art is now enabled by perform a great variety of assays which will detect the modulation of CalDAG-GEF or cAMP-GEF activity by candidate compounds. Particularly preferred and contemplated embodiments are discussed in some detail below.

A. CalDAG-GEF and/or cAMP-GEF Expression

In one series of embodiments, specific measures of CalDAG-GEF or cAMP-GEF expression are employed to screen candidate compounds for their ability to affect CalDAG-GEF or cAMP-GEF activity. Thus, using the CalDAG-GEF or cAMP-GEF nucleic acids and antibodies disclosed and otherwise enabled herein, one may use mRNA levels or protein levels as a marker for the ability of a candidate compound to modulate CalDAG-GEF or cAMP-GEF activity. The use of such probes and antibodies to measure gene and protein expression is well known in the art and discussed elsewhere herein.

B. Intracellular Localization

In another series of embodiments, compounds may be screened for their ability to modulate the activity of the CalDAG-GEFs or cAMP-GEFs based upon their effects on the

trafficking and intracellular localization of the CalDAG-GEFs or cAMP-GEFs. Differences in localization of mutant and normal CalDAG-GEFs and/or cAMP-GEFs may contribute to the etiology of CalDAG-GEF and/or cAMP-GEF-associated diseases. Compounds which can affect the localization of the CalDAG-GEFs and/or cAMP-GEFs may, therefore, be identified as potential therapeutics. Standard techniques known in the art may be employed to detect the localization of the CalDAG-GEFs and/or cAMP-GEFs. Generally, these techniques will employ the antibodies of the present invention, and in particular antibodies which selectively bind to one or more mutant CalDAG-GEFs or cAMP-GEFs but not to normal CalDAG-GEFs or cAMP-GEFs. As is well known in the art, such antibodies may be labeled by any of a variety of techniques (*e.g.*, fluorescent or radioactive tags, labeled secondary antibodies, avidin-biotin, etc.) to aid in visualizing the intracellular location of the CalDAG-GEFs or cAMP-GEFs. The CalDAG-GEFs or cAMP-GEFs may be co-localized to particular structures, as is known in the art, using antibodies to markers of those structures (*e.g.*, TGN38 for the Golgi, transferrin receptor for post-Golgi transport vesicles, LAMP2 for lysosomes). Western blots of purified fractions from cell lysates enriched for different intracellular membrane bound organelles (*e.g.*, lysosomes, synaptosomes, Golgi) may also be employed. In addition, the relative orientation of different domains of the CalDAG-GEFs and/or cAMP-GEFs across cellular domains may be assayed using, for example, electron microscopy and antibodies raised to those domains.

9. Screening and Diagnostics for CalDAG-GEF- or cAMP-GEF-associated disorders

A. General Diagnostic Methods

The CalDAG-GEF or cAMP-GEF genes and gene products, as well as the CalDAG-GEF or cAMP-GEF-derived probes, primers and antibodies, disclosed or otherwise enabled herein, are useful in the screening for carriers of alleles associated with CalDAG-GEF- or cAMP-GEF-associated disorders. Individuals at risk for such a disorder or individuals not previously known to be at risk, may be routinely screened using probes to detect the presence of a mutant CalDAG-GEF or cAMP-GEF gene or protein by a variety of techniques. Diagnosis of inherited cases of these diseases can be accomplished by methods based upon the nucleic acids (including genomic and mRNA/cDNA sequences), proteins, and/or antibodies disclosed and enabled herein, including functional assays designed to detect increases or decreases of the normal CalDAG-

GEF or cAMP-GEF activity and/or the presence of specific new activities conferred by the mutant CalDAG-GEFs or cAMP-GEFs. Preferably, the methods and products are based upon the human CalDAG-GEF or cAMP-GEF nucleic acids, proteins or antibodies, as disclosed or otherwise enabled herein. For brevity of exposition, but without limiting the scope of the invention, the following description will focus upon uses of the human homologues of CalDAG-GEF and cAMP-GEF. It will be understood, however, that homologous sequences from other species, including those disclosed herein, will be equivalent for many purposes.

B. Protein Based Screens and Diagnostics

When a diagnostic assay is to be based upon CalDAG-GEF or cAMP-GEF proteins, a variety of approaches are possible. For example, diagnosis can be achieved by monitoring differences in the electrophoretic mobility of normal and mutant proteins. Such an approach will be particularly useful in identifying mutants in which insertions, deletions or substitutions have resulted in a significant change in the electrophoretic migration of the resultant protein. Alternatively, diagnosis may be based upon differences in the proteolytic cleavage patterns of normal and mutant proteins, differences in molar ratios of the various amino acid residues, or by functional assays demonstrating altered function of the gene products.

C. Nucleic Acid Based Screens and Diagnostics

When the diagnostic assay is to be based upon nucleic acids from a sample, the assay may be based upon mRNA, cDNA or genomic DNA. Whether mRNA, cDNA, or genomic DNA is assayed, standard methods well known in the art may be used to detect the presence of a particular sequence either *in situ* or *in vitro* (See, e.g., Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL, 2nd ed. (1989)).

(1) Appropriate Probes and Primers

Whether for hybridization, RNase protection, ligase-mediated detection, PCR amplification or any other standards methods described herein and well known in the art, a variety of subsequences of the CalDAG-GEF and/or cAMP-GEF sequences disclosed or otherwise enabled herein will be useful as probes and/or primers. These sequences or subsequences will include both normal CalDAG-GEF or cAMP-GEF sequences and deleterious mutant sequences. In general, useful sequences will include at least 8-10, more preferably 10-15, and most preferably 15-25 consecutive nucleotides from the CalDAG-GEF or cAMP-GEF

introns, exons or intron/exon boundaries. Useful sequences will also include at least 25-500 consecutive nucleotides from the CalDAG-GEF or cAMP-GEF introns, exons or intron/exon boundaries. Depending upon the target sequence, the specificity required, and future technological developments, shorter sequences may also have utility. Therefore, any CalDAG-GEF or cAMP-GEF derived sequence which is employed to isolate, clone, amplify, identify or otherwise manipulate a CalDAG-GEF or cAMP-GEF sequence may be regarded as an appropriate probe or primer.

(2) Hybridization Screening

For *in situ* detection of a normal or mutant CalDAG-GEF, cAMP-GEF or other CalDAG-GEF and/or cAMP-GEF-associated nucleic acid sequence, a sample of tissue may be prepared by standard techniques and then contacted with one or more of the above-described probes, preferably one which is labeled to facilitate detection, and an assay for nucleic acid hybridization is conducted under stringent conditions which permit hybridization only between the probe and highly or perfectly complementary sequences.

(3) Restriction Mapping

Sequence alterations may also create or destroy fortuitous restriction enzyme recognition sites which are revealed by the use of appropriate enzyme digestion followed by electrophoresis and visualization. DNA fragments carrying the site (normal or mutant) are detected by their increase or reduction in size, or by the increase or decrease of corresponding restriction fragment numbers. Such restriction fragment length polymorphism analysis (RFLP), or restriction mapping, may be employed with genomic DNA, mRNA or cDNA. The CalDAG-GEF or cAMP-GEF sequences may be amplified by PCR using the above-described primers prior to restriction, in which case the lengths of the PCR products may indicate the presence or absence of particular restriction sites, and/or may be subjected to restriction after amplification. The CalDAG-GEF or cAMP-GEF fragments may be visualized by any convenient means (*e.g.*, under UV light in the presence of ethidium bromide).

(4) PCR Mapping

In another series of embodiments, a single base substitution mutation may be detected based on differential PCR product length or production in PCR. Thus, primers which span mutant sites or which, preferably, have 3' termini at mutation sites, may be employed to amplify

a sample of genomic DNA, mRNA or cDNA from a subject. A mismatch at a mutational site may be expected to alter the ability of the normal or mutant primers to promote the polymerase reaction and, thereby, result in product profiles which differ between normal subjects and heterozygous and/or homozygous CalDAG-GEF or cAMP-GEF mutants.

5 (5) Electrophoretic Mobility

Genetic testing based on DNA sequence differences also may be achieved by detection of alterations in electrophoretic mobility of DNA, mRNA or cDNA fragments in gels. Small sequence deletions and insertions, for example, can be visualized by high resolution gel electrophoresis of single or double stranded DNA, or as changes in the migration pattern of DNA heteroduplexes in non-denaturing gel electrophoresis.

10 (6) Chemical Cleavage of Mismatches

Mutations in the CalDAG-GEFs or cAMP-GEFs may also be detected by employing the chemical cleavage of mismatch (CCM) method (*See, e.g., Saleeba et al., METHODS IN ENZYMOLOGY, 217: 286-295 (1993)*). In this technique, probes (up to ~ 1 kb) may be mixed with a sample of genomic DNA, cDNA or mRNA obtained from a subject. The sample and probes are mixed and subjected to conditions which allow for heteroduplex formation (if any). Preferably, both the probe and sample nucleic acids are double-stranded, or the probe and sample may be PCR amplified together, to ensure creation of all possible mismatch heteroduplexes. Mismatched T residues are reactive to osmium tetroxide and mismatched C residues are reactive to hydroxylamine. Because each mismatched A will be accompanied by a mismatched T, and each mismatched G will be accompanied by a mismatched C, any nucleotide differences between the probe and sample (including small insertions or deletions) will lead to the formation of at least one reactive heteroduplex. After treatment with osmium tetroxide and/or hydroxylamine to modify any mismatch sites, the mixture is subjected to chemical cleavage at any modified mismatch sites by, for example, reaction with piperidine. The mixture may then be analyzed by standard techniques such as gel electrophoresis to detect cleavage products which would indicate mismatches between the probe and sample.

25 (7) Other Methods

Various other methods of detecting CalDAG-GEF or cAMP-GEF mutations, based upon the CalDAG-GEF or cAMP-GEF sequences disclosed and otherwise enabled herein, will be

apparent to those of ordinary skill in the art. Any of these may be employed in accordance with the present invention. These include, but are not limited to, nuclease protection assays (S1 or ligase-mediated), ligated PCR, denaturing gradient gel electrophoresis (DGGE; *see, e.g.*, Fischer et al., 80 PROC. NAT'L ACAD. SCI (USA), 1578-83 (1983)), restriction endonuclease
5 fingerprinting combined with SSCP (REF-SSCP; *see, e.g.*, Liu et al., 18 BIOTECHNIQUES 470-79 (1995)), and the like.

D. Other Screens and Diagnostics

Diagnosis also can be made by observation of alterations in CalDAG-GEF or cAMP-GEF transcription, translation, and post-translational modification and processing as well as
10 alterations in the intracellular and extracellular trafficking of CalDAG-GEF or cAMP-GEF gene products in the brain and peripheral cells. Such changes will include alterations in the amount of CalDAG-GEF or cAMP-GEF messenger RNA and/or protein, alteration in phosphorylation state, abnormal intracellular location/distribution, abnormal extracellular distribution, etc. Such assays will include: Northern Blots (with CalDAG-GEF or cAMP-GEF-specific and non-specific
15 nucleotide probes), Western blots and enzyme-linked immunosorbent assays (ELISA) (with antibodies raised specifically to a CalDAG-GEF or a cAMP-GEF functional domain, including various post-translational modification states).

E. Screening and Diagnostic Kits

In accordance with the present invention, diagnostic kits are also provided which will
20 include the reagents necessary for the above-described diagnostic screens. For example, kits may be provided which include antibodies or sets of antibodies which are specific to one or more mutant epitopes. These antibodies may, in particular, be labeled by any of the standard means which facilitate visualization of binding. Alternatively, kits may be provided in which oligonucleotide probes or PCR primers, as described above, are present for the detection and/or
25 amplification of mutant CalDAG-GEF, cAMP-GEF or other CalDAG-GEF and/or cAMP-GEF-associated nucleotide sequences. Again, such probes may be labeled for easier detection of specific hybridization. As appropriate to the various diagnostic embodiments described above, the oligonucleotide probes or antibodies in such kits may be immobilized to substrates and appropriate controls may be provided.

10. Methods of Treatment

The present invention now provides a basis for therapeutic intervention in diseases which are associated to the CalDAG-GEFs or cAMP-GEFs in that they are caused, prevented, exacerbated, or alleviated, or which may be caused, prevented, exacerbated, or alleviated, by the
5 either normal or mutant CalDAG-GEFs or cAMP-GEFs. In considering the various therapies described below, it is understood that such therapies may be targeted at tissue other than the brain where CalDAG-GEF or cAMP-GEF are also expressed.

Therapies to treat CalDAG-GEF and/or cAMP-GEF-associated diseases may be based upon (1) administration of normal CalDAG-GEF or cAMP-GEF proteins, (2) gene therapy with
10 normal CalDAG-GEF or cAMP-GEF genes to compensate for or replace the mutant genes, (3) gene therapy based upon antisense sequences to mutant CalDAG-GEF or cAMP-GEF genes or which "knock-out" the mutant genes, (4) gene therapy based upon sequences which encode a protein which blocks or corrects the deleterious effects of CalDAG-GEF or cAMP-GEF mutants, (5) immunotherapy based upon antibodies to normal and/or mutant CalDAG-GEF or cAMP-GEF
15 proteins, or (6) small molecules (drugs) which alter CalDAG-GEF or cAMP-GEF expression, block abnormal interactions between mutant forms of CalDAG-GEF or cAMP-GEF and other proteins or ligands, or which otherwise block the aberrant function of mutant CalDAG-GEF or cAMP-GEF proteins by altering the structure of the mutant proteins, by enhancing their metabolic clearance, or by inhibiting their function.

20 A. Protein Therapy

Treatment of CalDAG-GEF and/or cAMP-GEF-associated disorders, or disorders resulting from CalDAG-GEF and/or cAMP-GEF mutations, may be performed by providing an excess of inactive mutant protein to decrease the effect of the normal function of the protein, or by providing an excess of normal protein to reduce the effect of any aberrant function of the
25 mutant protein, by replacing a mutant protein with normal protein, or by modulating the function of the mutant protein.

B. Gene Therapy

In one series of embodiments, gene therapy may be employed in which normal or mutant copies of the CalDAG-GEF gene or the cAMP-GEF gene are introduced into patients to code
30 successfully for normal or mutant protein in one or more different affected cell types. The gene

must be delivered to those cells in a form in which it can be taken up and code for sufficient protein to provide effective function. Thus, it is preferred that the recombinant gene be operably joined to a strong promoter so as to provide a high level of expression which will compensate for, or out-compete, the naturally-occurring proteins. As noted above, the recombinant construct
5 may contain endogenous or exogenous regulatory elements, inducible or repressible regulatory elements, or tissue-specific regulatory elements.

In another series of embodiments, gene therapy may be employed to replace the naturally-occurring gene by homologous recombination with a recombinant construct. The recombinant construct may contain a normal or a mutant copy of the targeted CalDAG-GEF
10 and/or cAMP-GEF gene, in which case the defect is corrected *in situ*, or may contain a "knock-out" construct which introduces a stop codon, missense mutation, or deletion which abolished function of the mutant gene. It should be noted in this respect that such a construct may knock-out both the normal and mutant copies of the targeted CalDAG-GEF and/or cAMP-GEF gene in a heterozygous individual, but the total loss of CalDAG-GEF and/or cAMP-GEF gene function
15 may be less deleterious to the individual than continued progression of the disease state.

In another series of embodiments, antisense gene therapy may be employed. The antisense therapy is based on the fact that sequence-specific suppression of gene expression can be achieved by intracellular hybridization between mRNA or DNA and a complementary antisense species. The formation of a hybrid duplex may then interfere with the transcription of
20 the gene and/or the processing, transport, translation and/or stability of the target CalDAG-GEF and/or cAMP-GEF mRNA. Antisense strategies may use a variety of approaches including the administration of antisense oligonucleotides or antisense oligonucleotide analogs (*e.g.*, analogs with phosphorothioate backbones) or transfection with antisense RNA expression vectors. Again, such vectors may include exogenous or endogenous regulatory regions, inducible or
25 repressible regulatory elements, or tissue-specific regulatory elements.

In another series of embodiments, gene therapy may be used to introduce a recombinant construct encoding a protein or peptide which blocks or otherwise corrects the aberrant function caused by a naturally-occurring CalDAG-GEF and/or cAMP-GEF gene. In one embodiment, the recombinant gene may encode a peptide which corresponds to a mutant domain of a CalDAG-
30 GEF and/or cAMP-GEF which has been found to abnormally interact with another cell protein or

other cell ligand. Alternatively, the portion of a protein which interacts with a mutant, but not a normal, CalDAG-GEF and/or cAMP-GEF may be encoded and expressed by a recombinant construct in order to compete with, and thereby inhibit or block, the aberrant interaction.

5 Retroviral vectors can be used for somatic cell gene therapy especially because of their high efficiency of infection and stable integration and expression. The targeted cells however must be able to divide and the expression of the levels of normal protein should be high. The full length CalDAG-GEF or cAMP-GEF genes, subsequences encoding functional domains of the CalDAG-GEFs or cAMP-GEFs, or any of the other therapeutic peptides described above, can be cloned into a retroviral vector and driven from its endogenous promoter, from the retroviral long
10 terminal repeat, or from a promoter specific for the target cell type of interest. Other viral vectors which can be used include adeno-associated virus, vaccinia virus, bovine papilloma virus, or a herpes virus such as Epstein-Barr virus.

C. Immunotherapy

15 Antibodies may be raised to a mutant CalDAG-GEF or cAMP-GEF protein (or a portion thereof) and be administered to a patient to bind or block the mutant protein and prevent its deleterious effects. Alternatively, antibodies may be raised to specific complexes between mutant or wild-type CalDAG-GEF or cAMP-GEF and their interaction partners.

A further approach is to stimulate endogenous antibody production to the desired antigen. An immunogenic composition may be prepared as injectables, as liquid solutions or emulsions.
20 The CalDAG-GEF or cAMP-GEF protein or other antigen may be mixed with pharmaceutically acceptable excipients compatible with the protein. Such excipients may include water, saline, dextrose, glycerol, ethanol and combinations thereof. The immunogenic composition and vaccine may further contain auxiliary substances such as emulsifying agents or adjuvants to enhance effectiveness. Immunogenic compositions and vaccines may be administered
25 parenterally by injection subcutaneously or intramuscularly.

The immunogenic preparations and vaccines are administered in such amount as will be therapeutically effective, protective and immunogenic. Dosage depends on the route of administration and will vary according to the size of the host.

D. Small Molecule Therapeutics

As described and enabled herein, the present invention provides for a number of methods of identifying small molecules or other compounds which may be useful in the treatment of CalDAG-GEF- or cAMP-GEF-associated disorders. Thus, for example, the present invention provides for methods of identifying CalDAG-GEF or cAMP-GEF binding proteins and, in particular, methods for identifying proteins or other cell components which bind to or otherwise interact with mutant CalDAG-GEFs or cAMP-GEFs but not with the normal CalDAG-GEFs or cAMP-GEFs. The invention also provides for methods of identifying small molecules which can be used to disrupt undesired interactions between CalDAG-GEFs or cAMP-GEFs and other proteins or other cell components.

By identifying these proteins and analyzing these interactions, it is possible to screen for or design compounds which counteract or prevent the interaction, thereby, providing treatment for abnormal interactions. Therapies can be designed to modulate these interactions and thereby, to modulate CalDAG-GEF- or cAMP-GEF-associated disorders. The potential efficacy of these therapies can be tested by analyzing the affinity and function of these interactions after exposure to the therapeutic agent by standard pharmacokinetic measurements of affinity (*e.g.*, K_d , V_{max}) using synthetic peptides or recombinant proteins corresponding to functional domains of the CalDAG-GEF gene, the cAMP-GEF gene or other CalDAG-GEF and/or cAMP-GEF homologues. Another method for assaying the effect of any interactions involving functional domains is to monitor changes in the intracellular trafficking and post-translational modification of the relevant genes by *in situ* hybridization, immunohistochemistry, Western blotting and metabolic pulse-chase labeling studies in the presence of, and in the absence of, the therapeutic agents. A further method is to monitor the effects of "downstream" events including changes in second messenger events, *e.g.*, cAMP, intracellular Ca^{2+} , protein kinase activities, etc.

The effect of potential therapeutic agents in cell lines and whole animals can be monitored by monitoring transcription, translation, and post-translational modification of the CalDAG-GEF and/or cAMP-GEF proteins. Methods for these studies include Western and Northern blots, immunoprecipitation after metabolic labelling (pulse-chase) with radio-labelled methionine and ATP, and immunohistochemistry. The effect of these agents can also be monitored using studies which examine the relative binding affinities and relative amounts of

CalDAG-GEF or cAMP-GEF proteins involved in interactions with Rap1A, using either standard binding affinity assays or co-precipitation and Western blots using antibodies to Rap1A, CalDAG-GEF, cAMP-GEF, or other CalDAG-GEF and/or cAMP-GEF homologues.

Therapy using antisense oligonucleotides to block the expression of the mutant CalDAG-GEF gene or the mutant cAMP-GEF gene, co-ordinated with gene replacement with normal CalDAG-GEF or cAMP-GEF gene can also be applied using standard techniques of either gene therapy or protein replacement therapy.

V. Examples

Example 1: Isolation and characterization of CalDAG-GEF.

Human full-length CalDAG-GEFI cDNAs were isolated from a human frontal cortex λ ZAPII cDNA library (Stratagene) and a U937 λ ZAPII cDNA library. Mouse full-length CalDAG-GEFI was identified in the mouse EST database (GenBank accession number: W71787). Rat full-length CalDAG-GEFII cDNA was isolated from a rat whole brain λ ZAPII cDNA library by using human CalDAG-GEFII as a probe. Mouse ESTs identified through BLAST searches were purchased from Genome Systems Inc. (St. Louis, MO).

CalDAG-GEFI encodes an approximately 69-kD protein (Fig. 2D) that displays in its amino terminal region a GEF domain that is highly homologous to Ras-superfamily GEFs (Fig. 2A-2D). Multiple alignment analysis shows that genes of the CalDAG-GEF family form a cluster within the Ras-GEF superfamily distinct from Ras GEFs such as Sos1 and rRas-GEF (Fig. 2B). The region downstream of the GEF domain contains two tandem repeats of EF-hand Ca^{2+} binding motifs (Figs. 2A, 2E). The carboxy-terminal region displays a typical diacylglycerol/phorbol ester-binding domain, which is present in most PKC family proteins (Fig. 2A, 2F). Multiple sequence alignments and phylogenetic tree analysis were carried out with the LASERGENE Software Package (DNASTAR Inc.). Abbreviations and GenBank accession numbers of the protein sequences used in Figure 2 are as follows: C3G: 474982, mCdc25: 882120, rRas-GRF: 57665, hSos1 (human son-of-sevenless 1): 476780, BUD5: 171141, hCalmodulin: 115512, hCalbindin D28k: 227666, hCalcineurin B: 105504, hParvalbumin a: 131100, hTroponin C: 136043, hPKCa: 125549, hPKCb1: 125538, hPKCg: 462455.

To determine the small G protein target of CalDAG-GEFI, guanine nucleotide exchange activity *in vivo* was analyzed using intact 293T cells cotransfected with a eukaryotic expression construct of mouse CalDAG-GEFI and GST-tagged Ras family proteins. Full-length mouse CalDAG-GEFI cDNA inserted into a pCMV-SPORT expression vector with a carboxy-terminal
5 FLAG epitope was used for transfection. A PCR-amplified fragment of rat CalDAG-GEFII was subcloned into a pCAGGS expression vector with the addition of His₆-tag at its amino-terminus, resulting in pCAGGS-His-CalDAGII. pEBG-Krev1 that expresses Rap1A was used as a fusion protein to glutathione S-transferase (GST) in mammalian cells, as described in Gotoh et al., 15 Mol. Cell Biol. 6746-53 (1995), pEBG-R-Ras, other vectors for Ras-family proteins obtained by
10 inserting PCR-amplified cDNAs into pEBG expression vector, pCAGGS-C3G and pCAGGS-MSos1, and pCEV-H-RasV12. CalDAG-GEFI transfection produced a dramatic increase in GTP-bound Rap1A compared to the control but showed no or minimal activation of H-Ras, R-Ras, or Ral A. The increase in GTP-bound Rap1A was augmented in the presence of either the Ca²⁺ ionophore, A23187, or the phorbol ester, phorbol-12-myristate-13-acetate (TPA). Further,
15 A23187 and TPA had additive effects when administered together.

To determine the effect of CalDAG-GEFI on the Erk/MAP kinase cascade, Elk1 activation was measured in 293T cells transfected with CalDAG-GEFI or constitutively active H-Ras (RasV12), or both. 293T cells were transfected by SuperFect (Qiagen) as described in Gotoh, *supra*, with expression vectors for GST-tagged Ras family proteins and with those for
20 various GEFs. Cells were labeled 24 hours after transfection with ³²P_i for 2 hr. In some experiments, cells were stimulated with either 10 μM A23187 or 1 μM phorbol-12-myristate-13-acetate (TPA) for 3 min. GST-tagged Ras family proteins were collected from cell lysates with glutathione Sepharose. Guanine nucleotides bound to Ras family proteins were separated by thin layer chromatography (TLC). Activation of Elk1 was examined by the PathDetect Elk1
25 transreporting system (Stratagene). 293T cells were transfected with pFR-Luc and pFA-Elk1 with various expression vectors, and light output was detected and analyzed by the use of LAS1000 film. CalDAG-GEFI reduced RasV12 activation of Elk1 by approximately 4-fold and did not itself activate Elk1. Thus, CalDAG-GEFI strongly inhibits Ras-dependent stimulation of the Erk/MAP kinase cascade.

Northern analysis showed that human CalDAG-GEFI is expressed strongly in the brain and that CalDAG-GEFI mRNA is strikingly enriched in the striatum. Probes used included human CalDAG-GEFI: 729-bp EcoRI fragment, human CalDAG-GEFII: 584-bp SacI and HindIII fragment, rat CalDAG-GEFI: 439-bp fragment of EST clone RBC565 (GenBank accession number: C06861, and rat CalDAG-GEFII: 508-bp PCR amplified and subcloned fragment (nucleic acids 2541 to 3048 of SEQ ID NO:5). *In situ* hybridization of sections from the adult rat brain confirmed these restricted distribution patterns. Intense signal was present in the striatum (caudoputamen) and the ventral striatum (nucleus accumbens, olfactory tubercle). There was weaker signal in the olfactory bulb.

A series of monoclonal antibodies against the carboxy-terminal half of mouse CalDAG-GEFI were raised. His₆-tagged mouse CalDAG-GEFI polypeptide (amino acids 349 to 608 of SEQ ID NO:1) was expressed in bacteria, purified over Ni²⁺-nitrilotriacetic acid-agarose resin, and then used to immunize BALB/c mice. The resultant polyclonal antiserum was monitored by ELISA, Western blot, immunoprecipitation, and immunofluorescence assays on CalDAG-GEFI-transfected COS-7 cells. Hybridomas were generated by PEG (polyethylene glycol)-mediated fusion of donor splenocytes to the SP2/O cell line. Positive hybridoma cell lines were identified by screening in the assays described above, and purified by limiting dilution and single-cell cloning. Three hybridoma cell lines against mouse CalDAG-GEFI (mAbs 18B11, 2D9, and 18A7), in addition to the polyclonal fusion serum, were identified. Western analysis showed that mAbs 18B11 and 2D9 were specific for CalDAG-GEFI. Lightly post-fixed, cryostat-cut 10 μ m thick sections were immunostained by the ABC (Vectastain kit) method for CalDAG-GEFI with mAbs 18B11 and 2D9 and the polyclonal fusion serum, for tyrosine hydroxylase (TH) with monoclonal antibodies from INCSTAR, and for μ opioid receptor with polyclonal antiserum. Immunohistochemistry with mAb 18B11 showed a striking basal ganglia-enriched distribution pattern in sections of adult rat brain, with significant but weaker activity elsewhere. CalDAG-GEFI immunoreactivity marked the entire pathway from the striatal matrix compartment to the pallidum and substantia nigra pars reticulata, where very intense CalDAG-GEFI staining was present. Thus, CalDAG-GEFI is synthesized in striatal projection neurons and is transported to striatopallidal and striatonigral terminals.

To confirm that CalDAG-GEFI is synthesized in striatal projection neurons and transported to striatopallidal and striatonigral terminals in rats, intrastriatal injections of ibotenic acid (20 µg/µl, 1.5 µl per site, 5 day survival) were made unilaterally at 2 sites in the mid-lateral caudoputamen, with contralateral vehicle control injections were made. In other rats, unilateral subthalamic knife-cuts were made at an anteroposterior level between the entopeduncular nucleus and substantia nigra to sever the striatonigral efferents (1 and 3 days survivals), with control contralateral thalamic knife-cuts. These procedures all reduced CalDAG-GEFI staining in the substantia nigra. *In situ* hybridization was performed according to Simmons et al, 12 J. Histotechnol. 169-181 (1989). A 439bp rat EST clone RBC565 (98.4% identical to mouse CalDAG-GEFI nucleic acids 1777 to 2216 of SEQ ID NO:1) was isolated by BLAST search and used for making RNA probes with ³²P-labeled UTP (2,000 Ci/mmol, NEN, 1 Ci = 37 GBq) and T3 and T7 RNA polymerase. Brains were processed as above for CalDAG-GEFI and TH immunostaining. Thus, CalDAG-GEFI is a protein transported in striatal axons to their terminals. The terminal localization of CalDAG-GEFI was confirmed in subcellular fractionation experiments on dissected samples from the rat ventral midbrain, in which Western analysis showed the presence of CalDAG-GEFI in cytosol and in membrane fractions, including synaptosomes.

Because of the similarity of the GEF domains of CalDAG-GEFI and CalDAG-GEFII, the substrate specificity of CalDAG-GEFII with the same 293T cell assay system used for CalDAG-GEFI was examined. It was confirmed that CalDAG-GEFII activates Ras, and further shown that it activates H-Ras and R-Ras, but not Ral A or Rap1A. H-Ras activation was enhanced by A23187 and TPA. Moreover, CalDAG-GEFII, unlike CalDAG-GEFI, increased the transcriptional activity of Elk1 downstream to Erk/MAP kinase. Thus, in the 293T system, CalDAG-GEFI and CalDAG-GEFII target different Ras-superfamily small G proteins and have opposite effects on the MAP kinase cascade. Northern analysis further showed contrasting brain expression for CalDAG-GEFII, with highest expression being in the cerebellum, cerebral cortex, and amygdala, and low expression occurring in the striatum. Both genes are also expressed in hematopoietic organs in both human and rat.

Rap signaling is important in regulating basal ganglia output in response to Ca²⁺ and DAG. Corticostriatal inputs can activate the MAP kinase cascade in striatal projection neurons

(Sgambat et al., 18 J. Neurosci. 214-26 (1993)) and phosphoinositide (PI) signaling is strongly represented in these pathways (Fotuhi et al., 13 J. Neurosci. 3300-08 (1993)). Moreover, a number of receptor systems in the striatum and its striatonigral/striatopallidal pathways are linked to Ca^{2+} and PI signaling, notably including NMDA and metabotropic glutamate receptors, D_2 -class dopamine receptors, and tachykinin receptors (Fiorillo et al., 394 Nature 78-82 (1998)). A previously unrecognized signaling target for some of these systems is likely to be Rap1, via CalDAG-GEFI. In addition, CalDAG-GEFI has a synaptic function as demonstrated by the heavy accumulation of CalDAG-GEFI in the target nuclei of striatal outputs and the localization of Rap1 in synaptosomes and synaptic vesicles. The particular basal ganglia projection systems are enriched in CalDAG-GEFI and are differentially vulnerable to neurodegeneration in Huntington's disease.

Rap and Ras functions can be regulated coordinately or disjunctively by Ca^{2+} and DAG in the brain and hematopoietic organs, depending on the relative expression of CalDAG-GEFI and CalDAG-GEFII. In neurons, Ras/MAP kinase signaling has been directly implicated in synaptic transmission and the neuroplasticity underlying learning and memory. Different CalDAG-GEFI and CalDAG-GEFII expression patterns in the brain influence region-specific neuroplasticity mediated by Ca^{2+} and DAG signaling pathways. The presence of CalDAG-GEFI and CalDAG-GEFII in the hematopoietic system demonstrates the direct input of Ca^{2+} and DAG to Ras/Rap regulation of normal growth and differentiation as well as malignant transformation.

Example 2: Isolation and characterization of cAMP-GEFs.

cAMP-GEFI and cAMP-GEFII have similar domain structures, with a cAMP binding domain at the amino terminus and a GEF domain at the carboxy terminus separated by a link region (LR) (Fig. 3A). These mammalian proteins show strong structural homology to a predicted open reading frame (T20G5.5) in *C. elegans* cAMP-GEF (cel cAMP-GEF) (Fig. 3A). The cAMP binding domains of the cAMP-GEF family proteins form a distinct group within the cyclic nucleotide-binding protein superfamily and show the closest similarity to the B domains of PKA regulatory subunits (Fig. 3B). A PR(A/T)AT motif in the cAMP binding pocket is also conserved in the cAMP-GEF proteins (Fig. 3E). The first alanine of this motif confers cAMP

(alanine) as opposed to cGMP (threonine) binding specificity. All of the cAMP-GEF family members have alanine at this position, and therefore bind cAMP rather than cGMP.

The GEF domains of the cAMP-GEFs show high homology to those of Ras-GEF family proteins, but form an independent cluster distinct from Ras GEFs such as mCdc25, hSos1, and rRas-GRF (Fig. 3, C and D). The three structurally conserved regions specific to Ras-GEF family proteins (SCR1, SCR2, and SCR3) are present in all of the cAMP-GEF proteins (Fig. 3D). Multiple sequence alignments and phylogenetic tree analyses were carried out with LASERGENE (DNASTAR Inc.). Abbreviations and GenBank accession numbers of the protein sequences used in this figure: hPKARI α (human cAMP-dependent protein kinase regulatory subunit type I-alpha): 125193, hPKARI β : 1346362, hPKARII α : 125198, hPKARII β : 400115, hPKGII α (human cGMP-dependent protein kinase type I-alpha): 1255602, hPKGII β : 125379, hPKGII: 1906312, hCalDAG-GEFI: U71870, hCalDAG-GEFII: AF081195, C3G: 474982, hSos1 (human son-of-sevenless 1): 476780, mCdc25: 882120, rRas-GRF: 57665, BUD5: 171141.

In order to identify the small G protein substrate for cAMP-GEFI and II and the mode of cAMP regulation of GEF activity conferred by these proteins, the effects of cAMP-GEFI and cAMP-GEFII expression were analyzed in 293T cells on the ratio of GTP to GDP bound to different Ras family small G proteins in the presence or absence of forskolin and IBMX. Under basal conditions, in the absence of forskolin and IBMX, only Rap1 was activated significantly. In the presence of forskolin and IBMX, both cAMP-GEFI and II strongly and selectively activated Rap1A, but did not activate H-Ras, R-Ras or RalA. The effects of forskolin/IBMX treatment on cAMP-GEFI and II were dose-dependent with EC₅₀ values of 1.8 μ M and 0.3 μ M, respectively. Forskolin/IBMX treatment given alone had no effect.

A time-course analysis of the activation of Rap1A by forskolin/IBMX in cAMP-GEFI transfectants showed that the activation began within 10 sec, reached a maximum at 5 min, and continued for at least 60 min. Thus, cAMP-GEFI has a direct effect on Rap1A rather than secondary effects mediated by other Ras-superfamily GEFs. In addition, Sp-cAMPS, an analogue of cAMP, activated Rap1A at levels similar to those induced by forskolin/IBMX. Thus, cAMP has the capacity to activate the GEF domain of cAMP-GEFI.

Mutational analyses with cAMP-GEFI was performed to examine whether its cAMP-binding domain is required for the activation of Rap1A. In contrast to wild type cAMP-GEFI, a deletion mutant lacking a cAMP binding domain (pcDNA-rcAMP-GEFI:ΔcAMP(528) and (595)) did not activate Rap1A with or without forskolin/IBMX treatment. Mutants with a single
5 amino acid substitution at the cAMP binding pocket (pcDNA-rcAMP-GEFI:R(279)K) responded minimally to forskolin/IBMX treatment. Thus, the cAMP binding domain of cAMP-GEFI is necessary for its cAMP-dependent activation of Rap1A.

To assess further the cAMP binding capacity of cAMP-GEFI, a cAMP agarose affinity bead binding assay was performed. *In vitro* translated, radiolabeled cAMP-GEFI showed
10 selective binding to the beads that was competed by excess amounts of either cAMP or 8-Br-cAMP. cAMP-GEF protein can bind cAMP and that this binding can activate Rap1A.

cAMP-dependent activation of Rap1 has previously been ascribed to the phosphorylation of Rap1A by PKA, which raises its affinity to smgGDS, a GEF with broad substrate specificity. However, at least in the 293T cell assay system, an increase of GTP-bound Rap1A in response to
15 increasing cAMP levels with forskolin or treatment with the cAMP analogue, Sp-cAMPS was not detected in the absence of cAMP-GEFs. In addition, even in the presence of H-89, a potent and selective inhibitor of PKA, cAMP-GEFI and II could still activate Rap1A. The activation of Rap1A induced by cAMP-GEFI and II is independent of the PKA pathway.

Intracellular cAMP has been shown to interact directly with ion channels, but the vast
20 majority of cAMP-mediated effects in eukaryotes have been considered as sequels to cAMP binding by the regulatory subunits of the PKA tetramer. The diversity of physiological effects produced by cAMP have been attributed to the fact that, as a kinase, PKA has a large range of molecular targets. Reported herein are novel cAMP binding proteins that directly link the cAMP second messenger system to Ras superfamily signaling pathways and that appear selectively to
25 target Rap.

cAMP can exert profound cell-type specific effects on cell growth and differentiation and that cAMP is capable of inhibiting or stimulating the Ras/mitogen-activated protein (MAP) kinase/Erk pathway. The inhibition can occur at the initial translocation step by which Ras activates Raf, whereas activation of Rap1 is thought to occur through phosphorylation by PKA.
30 Activation of Rap1 has been suggested to be part of a switch mechanism determining whether

growth or differentiation occurs in response to nerve growth factor (NGF). cAMP-GEFs directly couple cAMP to Rap1, itself discovered as a negative regulator of Ras but suspected of having independent functions as well. Thus, different levels of cAMP-GEF expression confer cell-type specific regulation of Ras superfamily signaling systems.

5 The genes also exhibit developmentally regulated expression in the septum, medial thalamus and habenula, key structures in the limbic system variously linked to brain reward circuits, addiction and schizophrenia. Thus, cAMP-GEFs, in addition to PKA, underlie some of the neuronal functions of cAMP.

10 Example 3. Northern hybridization demonstrating the expression of CalDAG-GEFI and CalDAG-GEFII protein mRNAs in a variety of tissues.

15 Total cytoplasmic RNA was isolated from various human tissue samples including amygdala, cerebellum, corpus callosum, caudate nucleus, cortex, frontal lobe, hippocampus, kidney, liver, lung, medulla oblongata, occipital pole, putamen, spinal cord, substantia nigra, subthalamic nucleus, thalamus, and temporal lobe, obtained from surgical pathology using standard procedures such as CsCl purification. The RNA was then electrophoresed on a formaldehyde gel to permit size fractionation. The nitrocellulose membrane was prepared and the RNA was then transferred onto the membrane. ³²P-labeled cDNA probes were prepared and added to the membrane in order for hybridization between the probe the RNA to occur. After washing, the membrane was wrapped in plastic film and placed into imaging cassettes containing X-ray film. The autoradiographs were then allowed to develop for one to several days. Sizing was established by comparison to standard RNA markers. These northern blots demonstrated that the CalDAG-GEF genes are strongly expressed in the brain. Weaker hybridization was detectable elsewhere.

25 Example 4. Northern hybridization demonstrating the expression of cAMP-GEFI and cAMP-GEFII protein mRNAs in a variety of tissues.

30 Northern hybridization analysis was performed as in Example 3 to detect the expression of the cAMP-GEFI and cAMP-GEFII genes in a variety of human tissues. The tissues analyzed included adrenal gland, amygdala, bone marrow, cerebellum, corpus callosum, caudate nucleus,

colon (mucosal lining), caudputamen, cortex, frontal lobe, hippocampus, habenula, heart, kidney, liver, lung, lymph node, medulla obongata, occipital pole, olfactory bulb, ovary, pons, pancreas, putamen, septum, small intestines, skeletal muscle, spinal cord, spleen, stomach, substantia nigra, subthalamic nucleus, testis, thalamus, temporal lobe, thymus, trachea, and thyroid.

5 A striking contrast in the expression patterns of human cAMP-GEFI and II was observed by Northern analysis. Human cAMP-GEFI is widely expressed, with highest levels appearing in kidney, spleen, thyroid, heart, and pancreas. Human cAMP-GEFII shows a remarkably selective enrichment in the brain and the adrenal glands. Both genes were developmentally regulated. The expression patterns of the two genes in the nervous system also differ, with cAMP-GEFI
10 having a wider expression than cAMP-GEFII. These region-specific neuronal expression patterns were confirmed in *in situ* hybridization experiments. cAMP-GEFI mRNA is expressed broadly at low levels in the adult brain, but it is strongly and selectively expressed in parts of the neonatal brain, including the septum and the thalamus. By contrast, cAMP-GEFII is strongly expressed in the mature as well as the developing brain. Notable are the high levels of cAMP-
15 GEFII mRNA in the cerebral cortex, the hippocampus (especially CA3 and dentate gyrus), the habenula and the cerebellum. Genes of the cAMP-GEF family have widespread influence on cAMP functions in bodily organs and also region-specific functions in the brain.

Example 5. Isolation of CalDAG-GEF or cAMP-GEF binding proteins by yeast two-hybrid
20 system.

To identify proteins interacting with the CalDAG-GEF or cAMP-GEF proteins, a yeast expression plasmid vector (pAS2-1, Clontech) is generated by ligating an in-frame partial cDNA sequence encoding either residues of the CalDAG-GEF protein or residues of the cAMP-GEF protein into the EcoRI and BamHI sites of the vector. The resultant fusion protein contains the
25 GAL4 DNA binding domain coupled in-frame either to residues of the CalDAG-GEF protein or to residues of the cAMP-GEF protein. These expression plasmids are co-transformed, along with purified plasmid DNA from the human brain cDNA:pACT library, into yeast using the protocols of the Clontech Matchmaker yeast-two-hybrid kit (Clontech). Yeast clones bearing human brain cDNAs which interact with the CalDAG-GEF or cAMP-GEF fragments are
30 selected by HIS resistance and β gal⁺ activation. The clones are further selected by

cyclohexamide sensitivity and the inserts of the human brain cDNAs are isolated by PCR and sequenced.

Although preferred embodiments of the invention have been described herein in detail, it
5 will be understood by those skilled in the art that variations may be made thereto without
departing from the spirit of the invention or the scope of the following claims.

CLAIMS

What is claimed is:

1. An isolated nucleic acid comprising a nucleotide sequence encoding a protein selected from the group consisting of a normal CalDAG-GEFI protein, a mutant CalDAG-GEFI protein, a normal CalDAG-GEFII protein, and a mutant CalDAG-GEFII protein.

2. An isolated nucleic acid comprising a nucleotide sequence encoding a protein selected from the group consisting of a normal cAMP-GEFI protein, a mutant cAMP-GEFI protein, a normal cAMP-GEFII protein, and a mutant cAMP-GEFII protein.

3. An isolated nucleic acid as in claim 1 wherein said nucleic acid encodes a normal CalDAG-GEF protein and wherein said nucleotide sequence is selected from the group consisting of

(a) a sequence encoding a protein comprising the human CalDAG-GEFI amino acid sequence of SEQ ID NO: 4;

(b) a sequence encoding a protein comprising the murine CalDAG-GEFI amino acid sequence of SEQ ID NO: 2;

(c) a sequence encoding a protein comprising the human CalDAG-GEFII amino acid sequence of SEQ ID NO: 8; and

(d) a sequence encoding a protein comprising the murine CalDAG-GEFII amino acid sequence of SEQ ID NO: 6; and

(e) a sequence encoding a normal CalDAG-GEF protein and capable of hybridizing to a sequence complementary to any sequence of (a) - (d) under stringent hybridization conditions.

4. An isolated nucleic acid as in claim 2 wherein said nucleic acid encodes a normal cAMP-GEF protein and wherein said nucleotide sequence is selected from the group consisting of

(a) a sequence encoding a protein comprising the human cAMP-GEFI amino acid sequence of SEQ ID NO: 12;

(b) a sequence encoding a protein comprising the alternatively spliced human cAMP-GEFI amino acid sequence of SEQ ID NO: 14;

(c) a sequence encoding a protein comprising the rat cAMP-GEFI amino acid sequence of SEQ ID NO: 10;

(d) a sequence encoding a protein comprising the human cAMP-GEFII amino acid sequence of SEQ ID NO: 18;

(e) a sequence encoding a protein comprising the rat cAMP-GEFII amino acid sequence of SEQ ID NO: 16; and

(f) a sequence encoding a normal cAMP-GEF protein and capable of hybridizing to a sequence complementary to any sequence of (a) - (e) under stringent hybridization conditions.

5. An isolated nucleic acid comprising a nucleotide sequence of at least 8 consecutive nucleotides selected from the group consisting SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO. 17, and a sequence complementary to any of these sequences.

6. An isolated nucleic acid comprising a nucleotide sequence of at least 10 consecutive nucleotides selected from the group consisting SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO. 17, and a sequence complementary to any of these sequences.

7. An isolated nucleic acid comprising a nucleotide sequence of at least 15 consecutive nucleotides selected from the group consisting SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO. 17, and a sequence complementary to any of these sequences.

8. An isolated nucleic acid comprising a nucleotide sequence encoding at least one functional domain of a CalDAG-GEF protein selected from the group consisting of a normal CalDAG-GEFI protein, a mutant CalDAG-GEFI protein, a normal CalDAG-GEFII protein, and a mutant CalDAG-GEFII protein.

9. An isolated nucleic acid comprising a nucleotide sequence encoding at least one functional domain of a cAMP-GEF protein selected from the group consisting of a normal cAMP-GEFI protein, a normal cAMP-GEFII protein, a mutant cAMP-GEFI protein, and a mutant cAMP-GEFII protein.

10. An isolated nucleic acid comprising a nucleotide sequence encoding an antigenic determinant of a CalDAG-GEF protein selected from the group consisting of a normal CalDAG-GEFI protein, a normal CalDAG-GEFII protein, a mutant CalDAG-GEFI protein, and a mutant CalDAG-GEFII protein.

11. An isolated nucleic acid comprising a nucleotide sequence encoding an antigenic determinant of a cAMP-GEF protein selected from the group consisting of a normal cAMP-GEFI protein, a normal cAMP-GEFII protein, a mutant cAMP-GEFI protein, and a mutant cAMP-GEFII protein.

12. A method for identifying an allelic variant or heterospecific homologue of a human CalDAG-GEF gene comprising:
choosing a nucleic acid probe or primer capable of hybridizing to a human CalDAG-GEF gene sequence under stringent hybridization conditions;
mixing said probe or primer with a sample of nucleic acids which may contain a nucleic acid corresponding to said variant or homologue; and
detecting hybridization of said probe or primer to said nucleic acid corresponding to said variant or homologue.

1 13. A method as in claim 12 wherein said human CalDAG-GEF gene sequence is
2 selected from the group consisting of SEQ ID NO: 3 and SEQ ID NO: 7.

1 14. A method as in claim 12 wherein said sample comprises a sample of nucleic acids
2 selected from the group consisting of human genomic DNA, human mRNA, and human cDNA.

1 15. A method as in claim 12 wherein said sample comprises a sample of nucleic acids
2 selected from the group consisting of mammalian genomic DNA, mammalian mRNA, and
3 mammalian cDNA.

1 16. A method as in claim 12 wherein said sample comprises a sample of nucleic acids
2 selected from the group consisting of invertebrate genomic DNA, invertebrate mRNA, and
3 invertebrate cDNA.

1 17. A method as in claim 12 further comprising the step of isolating said nucleic acid
2 corresponding to said variant or homologue.

1 18. A method as in claim 12 wherein said nucleic acid is identified by hybridization.

1 19. A method as in claim 12 wherein said nucleic acid is identified by PCR amplification.

1 20. A method for identifying allelic variants or heterospecific homologues of a human
2 cAMP-GEF gene comprising:

3 choosing a nucleic acid probe or primer capable of hybridizing to a human cAMP-
4 GEF gene sequence under stringent hybridization conditions;

5 mixing said probe or primer with a sample of nucleic acids which may contain a
6 nucleic acid corresponding to said variant or homologue; and

7 detecting hybridization of said probe or primer to said nucleic acid corresponding to
8 said variant or homologue.

9

1 21. A method as in claim 12 wherein said human cAMP-GEF gene sequence is selected
2 from the group consisting of SEQ ID NO: 11, SEQ ID NO: 13, and SEQ ID NO: 17.

3

1 22. A method as in claim 20 wherein said sample comprises a sample of nucleic acids
2 selected from the group consisting of human genomic DNA, human mRNA, and human cDNA.

3

1 23. A method as in claim 20 wherein said sample comprises a sample of nucleic acids
2 selected from the group consisting of mammalian genomic DNA, mammalian mRNA, and
3 mammalian cDNA.

4

1 24. A method as in claim 20 wherein said sample comprises a sample of nucleic acids
2 selected from the group consisting of invertebrate genomic DNA, invertebrate mRNA, and
3 invertebrate cDNA.

4

1 25. A method as in claim 20 further comprising the step of isolating said nucleic acid
2 corresponding to said variant or homologue.

3

1 26. A method as in claim 20 wherein said nucleic acid is identified by hybridization.

2

1 27. A method as in claim 20 wherein said nucleic acid is identified by PCR amplification.

2

1 28. A method for identifying an allelic variant or heterospecific homologue of a human
2 CalDAG-GEF gene comprising:

3 choosing an antibody capable of selectively binding to a human CalDAG-GEF
4 protein;

5 mixing said antibody with a sample of proteins which may contain a protein
6 corresponding to said variant or homologue; and

detecting binding of said antibody to said protein corresponding to said variant or
homologue.

29. A method as in claim 28 wherein said sample comprises a sample of proteins selected
from the group consisting of human proteins, human fusion proteins, and proteolytic fragments
thereof.

30. A method as in claim 28 wherein said sample comprises a sample of nucleic acids
selected from the group consisting of mammalian proteins, mammalian fusion proteins, and
proteolytic fragments thereof.

31. A method as in claim 28 wherein said sample comprises a sample of nucleic acids
selected from the group consisting of invertebrate proteins, invertebrate fusion proteins, and
proteolytic fragments thereof.

32. A method as in claim 28 further comprising the step of substantially purifying said
protein corresponding to said variant or homologue.

33. A method for identifying an allelic variant or heterospecific homologue of a human
cAMP-GEF gene comprising:
choosing an antibody capable of selectively binding to a human cAMP-GEF protein;
mixing said antibody with a sample of proteins which may contain a protein
corresponding to said variant or homologue; and
detecting binding of said antibody to said protein corresponding to said variant or
homologue.

34. A method as in claim 33 wherein said sample comprises a sample of proteins selected
from the group consisting of human proteins, human fusion proteins, and proteolytic fragments
thereof.

4

1 35. A method as in claim 33 wherein said sample comprises a sample of proteins selected
2 from the group consisting of mammalian proteins, mammalian fusion proteins, and proteolytic
3 fragments thereof.

4

1 36. A method as in claim 33 wherein said sample comprises a sample of proteins selected
2 from the group consisting of invertebrate proteins, invertebrate fusion proteins, and proteolytic
3 fragments thereof.

4

1 37. A method as in claim 33 further comprising the step of substantially purifying said
2 protein corresponding to said variant or homologue.

3

1 38. An isolated nucleic acid comprising an allelic variant or a heterospecific homologue
2 of a gene selected from the group consisting of a human CalDAG-GEF gene, and a human
3 cAMP-GEF gene.

4

1 39. An isolated nucleic acid encoding an allelic variant or heterospecific homologue of a
2 protein selected from the group consisting of a human CalDAG-GEF protein, and a human
3 cAMP-GEF protein.

4

1 40. An isolated nucleic acid comprising a recombinant vector including a nucleotide
2 sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 5,
3 SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 13, SEQ ID NO: 15, SEQ ID NO:
4 17, and a sequence complementary to any of these sequences.

5

1 41. An isolated nucleic acid as in claim 40 wherein said vector is an expression vector
2 and said nucleotide sequence is operably joined to a regulatory region.

3

1 42. An isolated nucleic acid as in claim 41 wherein said expression vector may express
2 said nucleotide sequence in mammalian cells.
3

1 43. An isolated nucleic acid as in claim 42 wherein said cells are selected from the group
2 consisting of fibroblast, liver, kidney, spleen, bone marrow, and neurological cells.
3

1 44. An isolated nucleic acid as in claim 42 wherein said vector is selected from the group
2 consisting of vaccinia virus, adenovirus, retrovirus, neurotropic viruses, and Herpes simplex.
3

1 45. An isolated nucleic acid as in claim 41 wherein said expression vector encodes at
2 least a functional domain of a protein selected from the group consisting of normal CalDAG-
3 GEFI, a normal CalDAG-GEFII, a mutant CalDAG-GEFI, a mutant CalDAG-GEFII, a normal
4 cAMP-GEFI, a normal cAMP-GEFII, a mutant cAMP-GEFI, and a mutant cAMP-GEFII.
5

1 46. An isolated nucleic acid as in claim 41 wherein said vector further comprises
2 sequences encoding an exogenous protein operably joined to said nucleotide sequence and
3 whereby said vector encodes a fusion protein.
4

1 47. An isolated nucleic acid as in claim 46 wherein said exogenous protein is selected
2 from the group consisting of lacZ, trpE, maltose-binding protein, poly-His tags, and glutathione-
3 S-transferase.
4

1 48. An isolated nucleic acid comprising a recombinant expression vector including
2 nucleotide sequences corresponding to an endogenous regulatory region of a gene selected from
3 the group consisting of a CalDAG-GEF gene, and a cAMP-GEF gene.
4

1 49. An isolated nucleic acid as in claim 48 wherein said endogenous regulatory region is
2 operably joined to a marker gene.
3

1 50. A host cell transformed with an expression vector of any one of claims 41-49, or a
2 descendant thereof.
3

1 51. A host cell as in claim 50 wherein said host cell is selected from the group consisting
2 of bacterial cells and yeast cells.
3

1 52. A host cell as in claim 50 wherein said host cell is selected from the group consisting
2 of fetal cells, embryonic stem cells, zygotes, gametes, and germ line cells.
3

1 53. A host cell as in claim 50 wherein said cell is selected from the group consisting of
2 fibroblast, liver, kidney, spleen, bone marrow and neurological cells.
3

1 54. A host cell as in claim 50 wherein said cell is an invertebrate cell.
2

1 55. A non-human animal model for cancer, wherein a genome of said animal, or an
2 ancestor thereof, has been modified by at least one recombinant construct, and wherein said
3 recombinant construct has introduced a modification selected from the group consisting of

4 (a) insertion of nucleotide sequences encoding at least a functional domain of
5 a heterospecific normal CalDAG-GEF gene;

6 (b) insertion of nucleotide sequences encoding at least a functional domain of
7 a heterospecific mutant CalDAG-GEF gene;

8 (c) insertion of nucleotide sequences encoding at least a functional domain of
9 a conspecific homologue of a heterospecific mutant CalDAG-GEF gene;

10 (d) inactivation of an endogenous CalDAG-GEF gene;

11 (e) insertion of nucleotide sequences encoding at least a functional domain of
12 a heterospecific normal cAMP-GEF gene;

13 (f) insertion of nucleotide sequences encoding at least a functional domain of a
14 heterospecific mutant cAMP-GEF gene;

(g) insertion of nucleotide sequences encoding at least a functional domain of a conspecific homologue of a heterospecific mutant cAMP-GEF gene; and

(h) inactivation of an endogenous cAMP-GEF gene.

56. A non-human animal model as in claim 55 wherein said cancer is related to the Ras-pathway.

57. A non human animal model as in claim 56 wherein said cancer is selected from the group consisting of lung cancer, pancreatic cancer, breast cancer, colorectal cancer, and myeloid leukemia.

58. An animal model as in claim 55 wherein said modification is an insertion of a nucleotide sequence encoding at least a functional domain of a protein selected from the group consisting of a normal human CalDAG-GEF, and a normal cAMP-GEF gene.

59. An animal model as in claim 55 wherein said modification is an insertion of a nucleotide sequence encoding at least a functional domain of a protein selected from the group consisting of a mutant human CalDAG-GEF, and a mutant human cAMP-GEF gene.

60. An animal as in claim 55 wherein said animal is selected from the group consisting of rats, mice, hamsters, guinea pigs, rabbits, dogs, cats, goats, sheep, pigs, and non-human primates.

61. An animal as in claim 55 wherein said animal is an invertebrate.

62. A method for producing at least a functional domain of a protein selected from the group consisting of a CalDAG-GEF protein, and a cAMP-GEF protein, said method comprising culturing a host cell of any of claims 50-54 under suitable conditions to produce said protein by expressing said nucleic acid.

1 63. A substantially pure preparation of a protein selected from the group consisting of a
2 normal CalDAG-GEF protein, a mutant CalDAG-GEF protein, a normal cAMP-GEF protein,
3 and a mutant cAMP-GEF protein.
4

1 64. A substantially pure preparation as in claim 63 wherein said protein comprises a
2 normal protein selected from the group consisting of

- 3 (a) a protein comprising the amino acid sequence of SEQ ID NO: 2;
 - 4 (b) a protein comprising the amino acid sequence of SEQ ID NO: 4;
 - 5 (c) a protein comprising the amino acid sequence of SEQ ID NO: 6;
 - 6 (d) a protein comprising the amino acid sequence of SEQ ID NO: 8;
 - 7 (e) a protein comprising the amino acid sequence of SEQ ID NO: 10;
 - 8 (f) a protein comprising the amino acid sequence of SEQ ID NO: 12;
 - 9 (g) a protein comprising the amino acid sequence of SEQ ID NO: 14;
 - 10 (h) a protein comprising the amino acid sequence of SEQ ID NO: 16; and
 - 11 (i) a protein comprising the amino acid sequence of SEQ ID NO: 18.
- 12

1 65. A substantially pure preparation of a polypeptide comprising an amino acid sequence
2 of at least 4 consecutive amino acid residues selected from the group consisting of SEQ ID NO:
3 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID
4 NO: 14, SEQ ID NO: 16, and SEQ ID NO: 18.
5

1 66. A substantially pure preparation of a polypeptide comprising an amino acid sequence
2 of at least 10 consecutive amino acid residues selected from the group consisting of SEQ ID NO:
3 2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID
4 NO: 14, SEQ ID NO: 16, and SEQ ID NO: 18.
5

1 67. A substantially pure preparation of a polypeptide comprising an amino acid sequence
2 of at least 15 consecutive amino acid residues selected from the group consisting of SEQ ID NO:

2, SEQ ID NO: 4, SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, and SEQ ID NO: 18.

68. A substantially pure preparation of a polypeptide comprising at least one functional domain of a protein selected from the group consisting of a normal CalDAG-GEF protein, a mutant CalDAG-GEF protein, a normal cAMP-GEF protein, and a mutant cAMP-GEF protein.

69. A substantially pure preparation of a polypeptide comprising an antigenic determinant of a protein selected from the group consisting of a normal CalDAG-GEF protein, a mutant CalDAG-GEF protein, a normal cAMP-GEF protein, and a mutant cAMP-GEF protein.

70. A method of producing antibodies which selectively bind to a CalDAG-GEF protein comprising the steps of
administering an immunogenically effective amount of a CalDAG-GEF immunogen to an animal;
allowing said animal to produce antibodies to said immunogen; and
obtaining said antibodies from said animal or from a cell culture derived therefrom.

71. A method of producing antibodies which selectively bind to a cAMP-GEF protein comprising the steps of
administering an immunogenically effective amount of a cAMP-GEF immunogen to an animal;
allowing said animal to produce antibodies to said immunogen; and
obtaining said antibodies from said animal or from a cell culture derived therefrom.

72. A substantially pure preparation of an antibody which selectively binds to an antigenic determinant of a protein selected from the group consisting of a normal CalDAG-GEF protein, a mutant CalDAG-GEF protein, a normal cAMP-GEF protein, and a mutant cAMP-GEF protein.

5

1 73. A substantially pure preparation of an antibody as in claim 72 wherein said antibody
2 selectively binds to an antigenic determinant of a mutant CalDAG-GEF and fails to bind to a
3 normal CalDAG-GEF protein.

4

1 74. A substantially pure preparation of an antibody as in claim 72 wherein said antibody
2 selectively binds to an antigenic determinant of a mutant cAMP-GEF and fails to bind to a
3 normal cAMP-GEF protein.

4

1 75. A cell line producing an antibody of any one of claims 72-74.

2

1 76. A method for identifying compounds which can modulate the expression of a
2 CalDAG-GEF gene comprising:
3 contacting a cell with a test candidate wherein said cell includes a regulatory region of
4 a CalDAG-GEF gene operably joined to a coding region; and
5 detecting a change in expression of said coding region.

6

1 77. A method for identifying compounds which can modulate the expression of a cAMP-
2 GEF gene comprising:
3 contacting a cell with a test candidate wherein said cell includes a regulatory region of
4 a cAMP-GEF gene operably joined to a coding region; and
5 detecting a change in expression of said coding region.

6

1 78. A method as in claim 76 or 77 wherein said change comprises a change in a level of
2 an mRNA transcript encoded by said coding region.

3

1 79. A method as in claim 78 wherein said change comprises a change in a level of a
2 protein encoded by said coding region.

3

1 80. A method as in claim 78 wherein said change is a result of an activity of a protein
2 encoded by said coding region.
3

1 81. A method as in claim 78 wherein said coding region encodes a marker protein
2 selected from the group consisting of β -galactosidase, alkaline phosphatase, green fluorescent
3 protein, and luciferase.
4

1 82. A method for identifying compounds which can selectively bind to a CalDAG-GEF
2 protein comprising the steps of
3 providing a preparation including at least one CalDAG-GEF component;
4 contacting said preparation with a sample including at least one candidate compound;
5 and
6 detecting binding of said CalDAG-GEF component to said candidate compound.
7

1 83. A method for identifying compounds which can selectively bind to a cAMP-GEF
2 protein comprising the steps of
3 providing a preparation including at least one cAMP-GEF component;
4 contacting said preparation with a sample including at least one candidate compound;
5 and
6 detecting binding of said cAMP-GEF component to said candidate compound.
7

1 84. The method in claim 82 wherein said binding to said CalDAG-GEF component is
2 detected by an assay selected from the group consisting of: affinity chromatography, co-
3 immunoprecipitation, a Biomolecular Interaction Assay, and a yeast two-hybrid system.
4

1 85. The method in claim 83 wherein said binding to said cAMP-GEF component is
2 detected by an assay selected from the group consisting of: affinity chromatography, co-
3 immunoprecipitation, a Biomolecular Interaction Assay, and a yeast two-hybrid system.
4

1 86. A method of identifying compounds which can modulate activity of a CalDAG-GEF
2 comprising the steps of
3 providing a cell expressing a normal or mutant CalDAG-GEF gene;
4 contacting said cell with at least one candidate compound; and
5 detecting a change in a marker of said activity.
6

1 87. A method of identifying compounds which can modulate activity of a cAMP-GEF
2 comprising the steps of
3 providing a cell expressing a normal or mutant cAMP-GEF gene;
4 contacting said cell with at least one candidate compound; and
5 detecting a change in a marker of said activity.
6

1 88. A method as in claim 86 wherein measurement of said marker indicates a difference
2 between cells bearing an expressed mutant CalDAG-GEF gene and otherwise identical cells free
3 of an expressed mutant CalDAG-GEF gene.
4

1 89. A method as in claim 86 wherein said change comprises a change in a non-specific
2 marker of cell physiology selected from the group consisting of pH, intracellular calcium, cyclic
3 AMP levels, GTP/GDP ratios, phosphatidylinositol activity, and protein phosphorylation.
4

1 90. A method as in claim 86 wherein said change comprises a change in expression of
2 said CalDAG-GEF.
3

1 91. A method as in claim 86 wherein said change comprises a change in occurrence or
2 rate of apoptosis or cell death.
3

1 92. A method as in claim 86 wherein said cell is a cell cultured *in vitro*.
2

1 93. A method as in claim 92 wherein said cell is a transformed host cell of any one of
2 claims 50-54.

1 94. A method as in claim 92 wherein said cell is explanted from a host bearing at least
2 one mutant CalDAG-GEF gene.

1 95. A method as in claim 92 wherein said cell is explanted from a transgenic animal of
2 any one of claims 55-61.

1 96. A method as in claim 86 wherein said cell is a cell in a live animal.

1 97. A method as in claim 96 wherein said cell is a cell of a transgenic animal of any one
2 of claims 55-61.

1 98. A method as in claim 86 wherein said cell is in a human subject in a clinical trial.

1 99. A method as in claim 87 wherein measurement of said marker indicates a difference
2 between cells bearing an expressed mutant cAMP-GEF gene and otherwise identical cells free of
3 an expressed mutant cAMP-GEF gene.

1 100. A method as in claim 87 wherein said change comprises a change in a non-specific
2 marker of cell physiology selected from the group consisting of pH, intracellular calcium, cyclic
3 AMP levels, GTP/GDP ratios, phosphatidylinositol activity, and protein phosphorylation.

1 101. A method as in claim 87 wherein said change comprises a change in expression of
2 said cAMP-GEF.

1 102. A method as in claim 87 wherein said change comprises a change in occurrence or
2 rate of apoptosis or cell death.

3
1 103. A method as in claim 87 wherein said cell is a cell cultured *in vitro*.
2

1 104. A method as in claim 103 wherein said cell is a transformed host cell of any one of
2 claims 50-54.
3

1 105. A method as in claim 103 wherein said cell is explanted from a host bearing at least
2 one mutant cAMP-GEF gene.
3

1 106. A method as in claim 103 wherein said cell is explanted from a transgenic animal of
2 any one of claims 55-61.
3

1 107. A method as in claim 87 wherein said cell is a cell in a live animal.
2

1 108. A method as in claim 107 wherein said cell is a cell of a transgenic animal of any one
2 of claims 55-61.
3

1 109. A method as in claim 87 wherein said cell is in a human subject in a clinical trial.
2

1 110. A diagnostic method for determining if a subject bears a mutant CalDAG-GEF gene
2 comprising the steps of
3 providing a biological sample of said subject; and
4 detecting in said sample a mutant CalDAG-GEF nucleic acid, a mutant CalDAG-GEF
5 protein, or a mutant CalDAG-GEF activity.
6

1 111. A method as in claim 111, wherein a mutant CalDAG-GEF nucleic acid is detected
2 by an assay selected from the group consisting of direct nucleotide sequencing, probe specific
3 hybridization, restriction enzyme digest and mapping, PCR mapping, ligase-mediated PCR

detection, RNase protection, electrophoretic mobility shift detection, and chemical mismatch cleavage.

112. A method as in claim 110, wherein a mutant CalDAG-GEF protein is detected by an assay selected from the group consisting of an immunoassay, a protease assay, and an electrophoretic mobility assay.

113. A diagnostic method for determining if a subject bears a mutant cAMP-GEF gene comprising the steps of
providing a biological sample of said subject; and
detecting in said sample a mutant cAMP-GEF nucleic acid, a mutant cAMP-GEF protein, or a mutant cAMP-GEF activity.

114. A method as in claim 113, wherein a mutant cAMP-GEF nucleic acid is detected by an assay selected from the group consisting of direct nucleotide sequencing, probe specific hybridization, restriction enzyme digest and mapping, PCR mapping, ligase-mediated PCR detection, RNase protection, electrophoretic mobility shift detection, and chemical mismatch cleavage.

115. A method as in claim 113, wherein a mutant cAMP-GEF protein is detected by an assay selected from the group consisting of an immunoassay, a protease assay, and an electrophoretic mobility assay.

116. A pharmaceutical preparation comprising a substantially pure CalDAG-GEF protein and a pharmaceutically acceptable carrier.

117. A pharmaceutical preparation comprising a substantially pure cAMP-GEF protein and a pharmaceutically acceptable carrier.

1 118. A pharmaceutical preparation comprising an expression vector operably encoding a
2 protein selected from the group consisting of a CalDAG-GEF protein, and a cAMP-GEF protein,
3 wherein said expression vector may express said CalDAG-GEF protein or said cAMP-GEF
4 protein in a human subject, and a pharmaceutically acceptable carrier.

1 119. A pharmaceutical preparation comprising an expression vector operably encoding a
2 CalDAG-GEF antisense sequence, wherein said expression vector may express said CalDAG-
3 GEF antisense sequence in a human subject, and a pharmaceutically acceptable carrier.

1 120. A pharmaceutical preparation comprising an expression vector operably encoding a
2 cAMP-GEF antisense sequence, wherein said expression vector may express said cAMP-GEF
3 antisense sequence in a human subject, and a pharmaceutically acceptable carrier.

1 121. A pharmaceutical preparation comprising a substantially pure antibody, and a
2 pharmaceutically acceptable carrier,
3 wherein said antibody selectively binds to a mutant protein selected from the group
4 consisting of a mutant CalDAG-GEF protein, and a mutant cAMP-GEF protein.

1 122. A pharmaceutical preparation as in claim 121 wherein said preparation is essentially
2 free of an antibody which selectively binds a normal CalDAG-GEF protein.

1 123. A pharmaceutical preparation as in claim 121 wherein said preparation is essentially
2 free of an antibody which selectively binds a normal cAMP-GEF protein.

1 124. A pharmaceutical preparation comprising a substantially pure preparation of an
2 antigenic determinant of a mutant CalDAG-GEF protein or a mutant cAMP-GEF protein.

1 125. A pharmaceutical preparation as in claim 124 wherein said preparation is essentially
2 free of an antigenic determinant of a normal CalDAG-GEF protein.

GENES INTEGRATING SIGNAL TRANSDUCTION PATHWAYS

Abstract of the Disclosure

The present invention describes the identification, isolation, sequencing and characterization of two human CalDAG-GEF, and two human cAMP-GEF genes, which are associated with the Ras pathway. Also identified are CalDAG-GEF gene homologues in mice and cAMP-GEF gene homologues in rats. Nucleic acids and proteins comprising or derived from the CalDAG-GEFs and/or cAMP-GEFs are useful in screening and diagnosing certain Ras-associated cancers, in identifying and developing therapeutics for treatment of certain Ras-associated cancers, and in producing cell lines and transgenic animals useful as models of Ras-associated cancers.

479JAC5473/111.878507-1

479JAC5473/111.878507-1

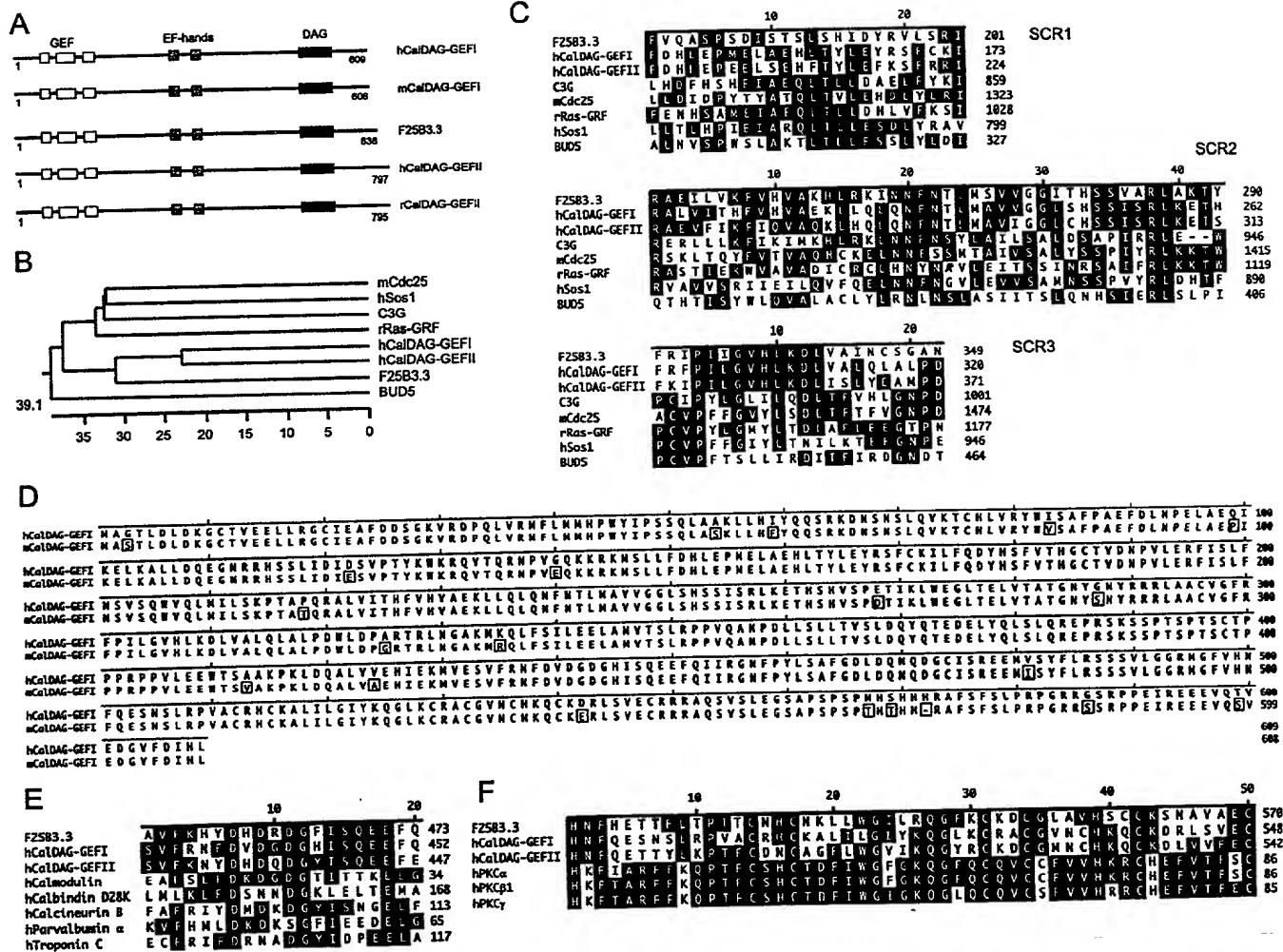
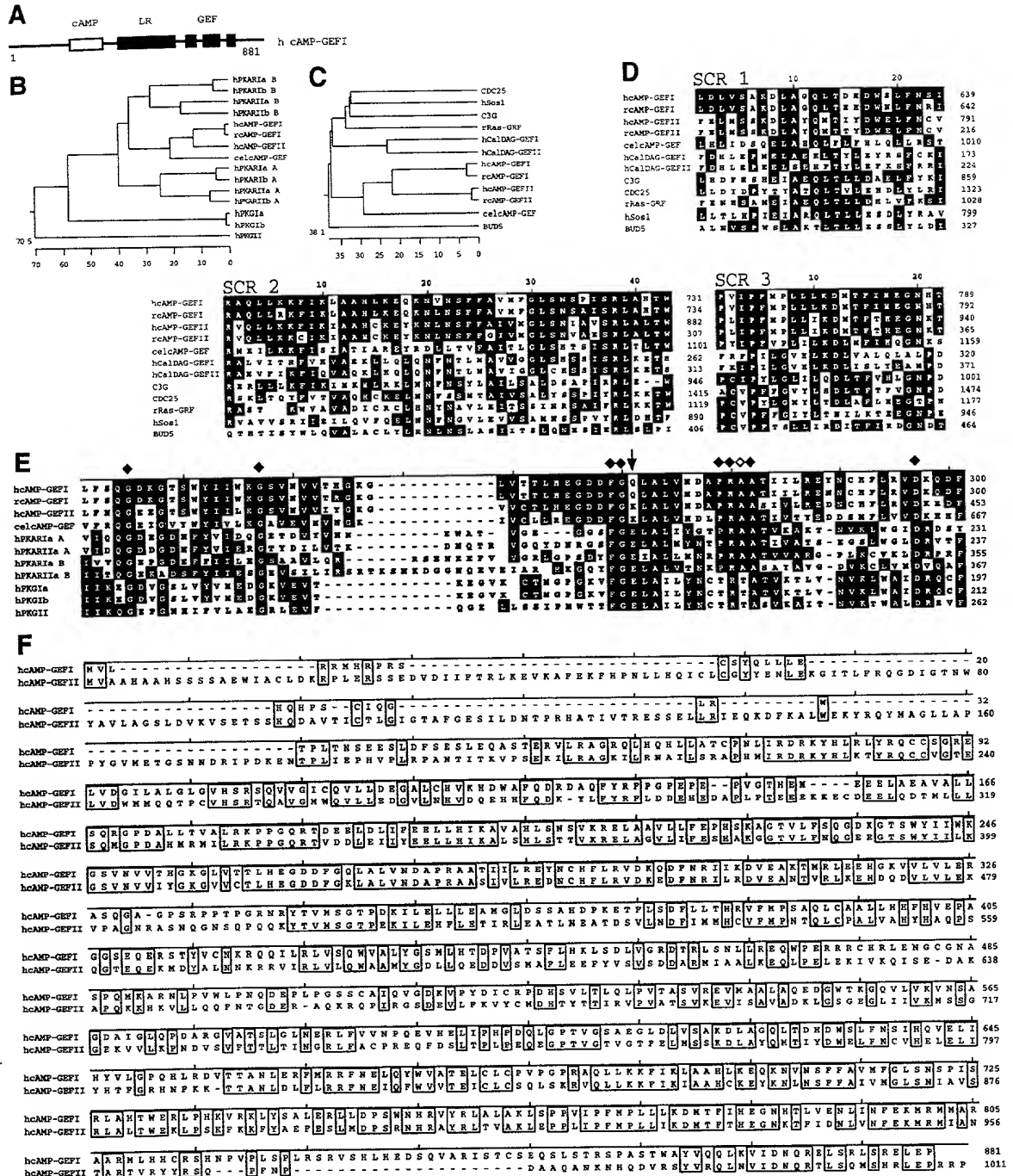


Figure 2

Figure 3



PATENT
Attorney Docket No. MIT-103

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANT(S): Kawasaki et al.

SERIAL NO.: Not yet assigned GROUP NO.: Not yet assigned

FILING DATE: October 22, 1999 EXAMINER: Not yet assigned

TITLE: Genes Integrating Signal Transduction Pathways

Box SEQUENCE

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

**SUBMISSION OF PAPER COPY AND/OR COMPUTER READABLE COPY
OF SEQUENCE LISTING FOR INVENTION
CONTAINING NUCLEOTIDE AND/OR AMINO ACID SEQUENCE**

(check and complete this item, if applicable)

1. ☐ This replies to the Patent Office Letter dated _____

Note: If these papers are filed before the office letter issues adequate identification of the original papers should be made, e.g., in addition to the name of the inventor and title of invention, the filing date based on the "Express Mail" procedure, the serial number from the return post card or the attorney's docket number added.

- ☐ A copy of the Patent Office Letter is enclosed.

2. Submitted herewith is/are

(check each item as applicable)

- A. ☐ a paper copy of the Sequence Listing for this application with each sequence assigned a separate identifier.
- B. ☒ a copy, in computer readable form, of the Sequence Listing for this application.

STATEMENT

3. I hereby state that:

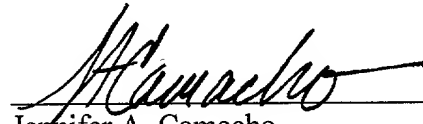
(complete applicable items A, B and/or C)

- A. ☐ the content of the paper and computer readable copies submitted herewith are the same.
- B. ☒ the content of the computer readable copy submitted herewith is the same as the Sequence Listing appearing on pages 60 to 108 of the original specification as filed on even date herewith.
- C. ☒ this submission includes no new matter.

Respectfully submitted,

Date: 10/22/99
Reg. No. 43,526

Tel. No.: (617) 248-7476
Fax No.: (617) 248-7100



Jennifer A. Camacho
Attorney for Applicant(s)
Testa, Hurwitz, & Thibault, LLP
High Street Tower
125 High Street
Boston, Massachusetts 02110

479JAC5473/111.A877137-1

SEQUENCE LISTING

<110> Kawasaki, Hiroaki
Graybiel, Ann
Housman, David

<120> Genes Integrating Signal Transduction Pathways

<130> MIT-103

<140>
<141>

<150> US 60/105,507
<151> 1998-10-23

<150> US 60/108,685
<151> 1998-11-16

<160> 18

<170> PatentIn Ver. 2.0

<210> 1
<211> 2250
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (166)..(1989)
<223> CalDAG-GEFI

<400> 1
cgaaggatca gaggctgagc tggttcaagt gaacagaagg tctgggaggt gaactgcatt 60
cgggttttgca ttctgaagta aaggacttgg gacaggggta cgaatcgagc actgtgggag 120
gctctgagag tgtaacttgg gtctagccca ctggcaccgg cagcc atg gcg agc act 177
Met Ala Ser Thr
1
ctg gac ctg gac aag ggt tgc acc gtg gag gag ctg ctc cgt ggc tgt 225
Leu Asp Leu Asp Lys Gly Cys Thr Val Glu Glu Leu Leu Arg Gly Cys
5 10 15 20
atc gaa gcc ttt gat gac tct gga aag gtg cga gat cca cag cta gtg 273
Ile Glu Ala Phe Asp Asp Ser Gly Lys Val Arg Asp Pro Gln Leu Val
25 30 35
cgc atg ttt ctc atg atg cac ccc tgg tac ata cct tcc tct cag ctg 321
Arg Met Phe Leu Met Met His Pro Trp Tyr Ile Pro Ser Ser Gln Leu
40 45 50
gct tcg aaa ctg ctc cac ttc tat cag caa tcc cgg aag gac aac tcc 369
Ala Ser Lys Leu Leu His Phe Tyr Gln Gln Ser Arg Lys Asp Asn Ser
55 60 65
aat tcc cta cag gtg aaa acc tgt cac ctg gtc agg tac tgg gtc tca 417
Asn Ser Leu Gln Val Lys Thr Cys His Leu Val Arg Tyr Trp Val Ser
70 75 80
gcc ttc cca gca gag ttc gac ttg aac cca gag ctg gct gaa ccg atc 465
Ala Phe Pro Ala Glu Phe Asp Leu Asn Pro Glu Leu Ala Glu Pro Ile
85 90 95 100
aag gag ctg aag gct ctg tta gac caa gaa gga aac cgc agg cac agc 513
Lys Glu Leu Lys Ala Leu Leu Asp Gln Glu Gly Asn Arg Arg His Ser
105 110 115

agc ctc atc gac atc gag agt gtc ccc acc tac aag tgg aag cgg cag	561
Ser Leu Ile Asp Ile Glu Ser Val Pro Thr Tyr Lys Trp Lys Arg Gln	
120 125 130	
gtg acc cag cgg aac cct gtg gaa cag aaa aag cgc aag atg tcc ctg	609
Val Thr Gln Arg Asn Pro Val Glu Gln Lys Lys Arg Lys Met Ser Leu	
135 140 145	
ttg ttt gat cac ttg gag cct atg gaa ctg gca gaa cat ctc acc tac	657
Leu Phe Asp His Leu Glu Pro Met Glu Leu Ala Glu His Leu Thr Tyr	
150 155 160	
ttg gag tat cgg tcc ttc tgc aag atc ctg ttc cag gac tat cac agc	705
Leu Glu Tyr Arg Ser Phe Cys Lys Ile Leu Phe Gln Asp Tyr His Ser	
165 170 175 180	
ttt gtg act cat ggc tgc act gta gac aat ccg gtc ctg gag cga ttc	753
Phe Val Thr His Gly Cys Thr Val Asp Pro Val Leu Glu Arg Phe	
185 190 195	
atc tcc ctc ttc aac agt gtc tct cag tgg gtc caa ctc atg atc ctc	801
Ile Ser Leu Phe Asn Ser Val Ser Gln Trp Val Gln Leu Met Ile Leu	
200 205 210	
agc aag ccc aca gcc acg cag cgg gcg ctg gtc atc aca cat ttc gtg	849
Ser Lys Pro Thr Ala Thr Gln Arg Ala Leu Val Ile Thr His Phe Val	
215 220 225	
cat gtg gca gag aag ctg ctg cag ctg cag aac ttc aac acg ttg atg	897
His Val Ala Glu Lys Leu Gln Leu Gln Asn Phe Asn Thr Leu Met	
230 235 240	
gcc gtc gtg gga ggc ctg agc cac agc tcc atc tca cgc ctc aag gag	945
Ala Val Val Gly Gly Leu Ser His Ser Ser Ile Ser Arg Leu Lys Glu	
245 250 255 260	
acc cac agc cat gtc agc cct gac acc atc aag ctc tgg gaa ggt ctg	993
Thr His Ser His Val Ser Pro Asp Thr Ile Lys Leu Trp Glu Gly Leu	
265 270 275	
aca gaa cta gtg aca gct act ggc aac tac agc aac tac cgg cga agg	1041
Thr Glu Leu Val Thr Ala Thr Gly Asn Tyr Ser Asn Tyr Arg Arg Arg	
280 285 290	
ctg gcg gcc tgc gtg ggc ttc cgc ttt cct atc ctg ggt gtg cac ctc	1089
Leu Ala Ala Cys Val Gly Phe Arg Phe Pro Ile Leu Gly Val His Leu	
295 300 305	
aag gat cta gtg gct ctg cag ctg gct ctg cct gac tgg ctg gac cca	1137
Lys Asp Leu Val Ala Leu Gln Leu Ala Leu Pro Asp Trp Leu Asp Pro	
310 315 320	
ggt cgg acc cgg ctc aat gga gcc aag atg agg cag ctt ttc agc att	1185
Gly Arg Thr Arg Leu Asn Gly Ala Lys Met Arg Gln Leu Phe Ser Ile	
325 330 335 340	
ctg gag gag ttg gct atg gtg acc agt ctt cga cca cca gtg caa gcc	1233
Leu Glu Glu Leu Ala Met Val Thr Ser Leu Arg Pro Pro Val Gln Ala	
345 350 355	
aac ccc gac ctg ctg agt ctg ctc acg gtg tcc ctg gac cag tat cag	1281
Asn Pro Asp Leu Leu Ser Leu Leu Thr Val Ser Leu Asp Gln Tyr Gln	
360 365 370	
acg gag gat gag ctg tat cag ctc tct ctg cag cga gag cca cgt tcc	1329
Thr Glu Asp Glu Leu Tyr Gln Leu Ser Leu Gln Arg Glu Pro Arg Ser	
375 380 385	

120 125 130 135 140 145 150 155 160 165 170 175 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355 360 365 370 375 380 385

aag tca tcg ccc acc agc ccc acc agc tgc acc ccg cct ccc cgg ccg	1377
Lys Ser Ser Pro Thr Ser Pro Thr Ser Cys Thr Pro Pro Pro Arg Pro	
390 395 400	
cct gtg ctg gaa gag tgg acc tca gtt gcc aag cct aag ctg gac caa	1425
Pro Val Leu Glu Glu Trp Thr Ser Val Ala Lys Pro Lys Leu Asp Gln	
405 410 415 420	
gcc ttg gtg gcc gag cac att gag aag atg gtg gag tct gtg ttc cgg	1473
Ala Leu Val Ala Glu His Ile Glu Lys Met Val Glu Ser Val Phe Arg	
425 430 435	
aac ttt gac gtt gat ggg gac ggt cac atc tcc cag gag gag ttc cag	1521
Asn Phe Asp Val Asp Gly Asp Gly His Ile Ser Gln Glu Glu Phe Gln	
440 445 450	
atc atc cgg ggc aac ttc cct tat ctc agc gcc ttt ggg gac ttg gac	1569
Ile Ile Arg Gly Asn Phe Pro Tyr Leu Ser Ala Phe Gly Asp Leu Asp	
455 460 465	
cag aac cag gat ggc tgc atc agc cgg gag gag atg att tcc tac ttc	1617
Gln Asn Gln Asp Gly Cys Ile Ser Arg Glu Glu Met Ile Ser Tyr Phe	
470 475 480	
ctg cgc tcc agc tcc gtg ctg gga ggc cgc atg ggc ttc gta cac aac	1665
Leu Arg Ser Ser Ser Val Leu Gly Gly Arg Met Gly Phe Val His Asn	
485 490 495 500	
ttc cag gag agt aac tcg cta agg ccg gtc gcc tgc cga cac tgc aaa	1713
Phe Gln Glu Ser Asn Ser Leu Arg Pro Val Ala Cys Arg His Cys Lys	
505 510 515	
gct ctg atc ctg ggc atc tac aag cag ggc ctc aaa tgt aga gct tgt	1761
Ala Leu Ile Leu Gly Ile Tyr Lys Gln Gly Leu Lys Cys Arg Ala Cys	
520 525 530	
ggt gtg aac tgc cac aag cag tgc aaa gac cga ctg tca gtg gaa tgt	1809
Gly Val Asn Cys His Lys Gln Cys Lys Asp Arg Leu Ser Val Glu Cys	
535 540 545	
cgc cgc cgc gcc cag agt gtg agc ctg gag ggc tct gca ccc tct ccc	1857
Arg Arg Arg Ala Gln Ser Val Ser Leu Glu Gly Ser Ala Pro Ser Pro	
550 555 560	
tca ccc aca cat acc cac cat cgg gcc ttc agc ttc tcc ctg cct cgc	1905
Ser Pro Thr His Thr His His Arg Ala Phe Ser Phe Ser Leu Pro Arg	
565 570 575 580	
cca ggc agg cgc agc tct cgg cct cca gag atc cgt gaa gag gag gtg	1953
Pro Gly Arg Arg Ser Ser Arg Pro Pro Glu Ile Arg Glu Glu Glu Val	
585 590 595	
cag act gtg gaa gat ggt gtg ttc gac atc cac tta taagacgtg	1999
Gln Thr Val Glu Asp Gly Val Phe Asp Ile His Leu	
600 605	
tgactatcaa ggactcattc ctgccttgga gaaaagactt ggagcagagc agggagccag	2059
ggattctggg gcaggaggtt ggggctgaag gtgggggaag ttgaagggtg catgcactga	2119
aaaaaaggcc agggctggtg tccctaaggt tgtacagact tctgtgaata tttgtatttt	2179
ccagatggaa taaaaaggcc cgaataatta acctcgaaaa aaaaaaaaaa aaaaaaaaaa	2239
aaaaaaaaaa a	2250

<210> 2
<211> 608

<212> PRT
<213> Mus musculus

<400> 2
Met Ala Ser Thr Leu₅ Asp Leu Asp Lys Gly₁₀ Cys Thr Val Glu Glu₁₅ Leu
1
Leu Arg Gly Cys₂₀ Ile Glu Ala Phe Asp₂₅ Asp Ser Gly Lys Val Arg Asp
Pro Gln Leu Val Arg Met Phe₄₀ Leu Met Met His Pro Trp₄₅ Tyr Ile Pro
Ser Ser₅₀ Gln Leu Ala Ser Lys₅₅ Leu Leu His Phe Tyr₆₀ Gln Gln Ser Arg
Lys Asp Asn Ser Asn Ser₇₀ Leu Gln Val Lys Thr₇₅ Cys His Leu Val Arg₈₀
Tyr Trp Val Ser Ala₈₅ Phe Pro Ala Glu Phe₉₀ Asp Leu Asn Pro Glu Leu₉₅
Ala Glu Pro Ile₁₀₀ Lys Glu Leu Lys Ala₁₀₅ Leu Leu Asp Gln Glu Gly Asn
Arg Arg His₁₁₅ Ser Ser Leu Ile Asp₁₂₀ Ile Glu Ser Val Pro Thr Tyr Lys
Trp Lys Arg Gln Val Thr Gln Arg Asn Pro Val Glu Gln Lys Lys Arg
130
Lys Met Ser Leu Leu Phe₁₅₀ Asp His Leu Glu Pro Met Glu Leu Ala Glu₁₆₀
His Leu Thr Tyr Leu₁₆₅ Glu Tyr Arg Ser Phe₁₇₀ Cys Lys Ile Leu Phe Gln₁₇₅
Asp Tyr His Ser₁₈₀ Phe Val Thr His Gly₁₈₅ Cys Thr Val Asp Asn Pro Val
Leu Glu Arg₁₉₅ Phe Ile Ser Leu Phe₂₀₀ Asn Ser Val Ser Gln Trp Val Gln
Leu Met Ile Leu Ser Lys Pro₂₁₅ Thr Ala Thr Gln Arg Ala Leu Val Ile
210
Thr His Phe Val His Val₂₃₀ Ala Glu Lys Leu Leu Gln Leu Gln Asn Phe₂₄₀
225
Asn Thr Leu Met Ala₂₄₅ Val Val Gly Gly₂₅₀ Leu Ser His Ser Ser Ile Ser₂₅₅
Arg Leu Lys Glu₂₆₀ Thr His Ser His Val₂₆₅ Ser Pro Asp Thr Ile Lys Leu
270
Trp Glu Gly Leu Thr Glu Leu Val₂₈₀ Thr Ala Thr Gly Asn Tyr Ser Asn
275
Tyr Arg Arg Arg Leu Ala Ala₂₉₅ Cys Val Gly Phe Arg₃₀₀ Phe Pro Ile Leu
290
Gly Val His Leu Lys Asp₃₁₀ Leu Val Ala Leu Gln Leu Ala Leu Pro Asp₃₂₀
305
Trp Leu Asp Pro Gly₃₂₅ Arg Thr Arg Leu Asn Gly Ala Lys Met Arg Gln
330
Leu Phe Ser Ile₃₄₀ Leu Glu Glu Leu Ala₃₄₅ Met Val Thr Ser Leu Arg Pro
350

1000
900
800
700
600
500
400
300
200
100
0

Pro Val Gln Ala Asn Pro Asp Leu Leu Ser Leu Leu Thr Val Ser Leu
355 360 365
Asp Gln Tyr Gln Thr Glu Asp Glu Leu Tyr Gln Leu Ser Leu Gln Arg
370 375 380
Glu Pro Arg Ser Lys Ser Ser Pro Thr Ser Pro Thr Ser Cys Thr Pro
385 390 395 400
Pro Pro Arg Pro Pro Val Leu Glu Glu Trp Thr Ser Val Ala Lys Pro
405 410 415
Lys Leu Asp Gln Ala Leu Val Ala Glu His Ile Glu Lys Met Val Glu
420 425 430
Ser Val Phe Arg Asn Phe Asp Val Asp Gly Asp Gly His Ile Ser Gln
435 440 445
Glu Glu Phe Gln Ile Ile Arg Gly Asn Phe Pro Tyr Leu Ser Ala Phe
450 455 460
Gly Asp Leu Asp Gln Asn Gln Asp Gly Cys Ile Ser Arg Glu Glu Met
465 470 475 480
Ile Ser Tyr Phe Leu Arg Ser Ser Ser Val Leu Gly Gly Arg Met Gly
485 490 495
Phe Val His Asn Phe Gln Glu Ser Asn Ser Leu Arg Pro Val Ala Cys
500 505 510
Arg His Cys Lys Ala Leu Ile Leu Gly Ile Tyr Lys Gln Gly Leu Lys
515 520 525
Cys Arg Ala Cys Gly Val Asn Cys His Lys Gln Cys Lys Asp Arg Leu
530 535 540
Ser Val Glu Cys Arg Arg Arg Ala Gln Ser Val Ser Leu Glu Gly Ser
545 550 555 560
Ala Pro Ser Pro Ser Pro Thr His Thr His His Arg Ala Phe Ser Phe
565 570 575
Ser Leu Pro Arg Pro Gly Arg Arg Ser Ser Arg Pro Pro Glu Ile Arg
580 585 590
Glu Glu Glu Val Gln Thr Val Glu Asp Gly Val Phe Asp Ile His Leu
595 600 605

<210> 3
<211> 2236
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (161)..(1987)
<223> CalDAG-GEFI

<400> 3
ggggactcaa ggctggcctg gctcaagtga acagcacgtc caggaggcga cctcgctccgc 60
gggtttgcat tctggggtgg acgagctggg ggttcggtcc gagcccgggtg ggaggctccc 120
ggagcgcagc ctgggcccag cccaccccg cccggcgggc atg gca ggc acc ctg 175
Met Ala Gly Thr Leu
1 5

gac Asp	ctg Leu	gac Asp	aag Lys	ggc Gly 10	tgc Cys	acg Thr	gtg Val	gag Glu	gag Glu 15	ctg Leu	ctc Leu	cgc Arg	ggg Gly	tgc Cys 20	atc Ile	223
gaa Glu	gcc Ala	ttc Phe	gat Asp 25	gac Asp	tcc Ser	ggg Gly	aag Lys	gtg Val 30	cgg Arg	gac Asp	ccg Pro	cag Gln	ctg Leu 35	gtg Val	cgc Arg	271
atg Met	ttc Phe	ctc Leu 40	atg Met	atg Met	cac His	ccc Pro	tgg Trp 45	tac Tyr	atc Ile	ccc Pro	tcc Ser	tct Ser 50	cag Gln	ctg Leu	gcg Ala	319
gcc Ala	aag Lys 55	ctg Leu	ctc Leu	cac His	atc Ile	tac Tyr 60	caa Gln	caa Gln	tcc Ser	cgg Arg	aag Lys 65	gac Asp	aac Asn	tcc Ser	aat Asn	367
tcc Ser 70	ctg Leu	cag Gln	gtg Val	aaa Lys	acg Thr 75	tgc Cys	cac His	ctg Leu	gtc Val	agg Arg 80	tac Tyr	tgg Trp	atc Ile	tcc Ser	gcc Ala 85	415
ttc Phe	cca Pro	gcg Ala	gag Glu	ttt Phe 90	gac Asp	ttg Leu	aac Asn	ccg Pro	gag Glu 95	ttg Leu	gct Ala	gag Glu	cag Gln	atc Ile 100	aag Lys	463
gag Glu	ctg Leu	aag Lys	gct Ala 105	ctg Leu	cta Leu	gac Asp	caa Gln	gaa Glu 110	ggg Gly	aac Asn	cga Arg	cgg Arg	cac His 115	agc Ser	agc Ser	511
cta Leu	atc Ile	gac Asp 120	ata Ile	gac Asp	agc Ser	gtc Val	cct Pro 125	acc Thr	tac Tyr	aag Lys	tgg Trp	aag Lys 130	cgg Arg	cag Gln	gtg Val	559
act Thr	cag Gln 135	cgg Arg	aac Asn	cct Pro	gtg Val	gga Gly 140	cag Gln	aaa Lys	aag Lys	cgc Arg	aag Lys 145	atg Met	tcc Ser	ctg Leu	ttg Leu	607
ttt Phe 150	gac Asp	cac His	ctg Leu	gag Glu	ccc Pro 155	atg Met	gag Glu	ctg Leu	gcg Ala	gag Glu 160	cat His	ctc Leu	acc Thr	tac Tyr	ttg Leu 165	655
gag Glu	tat Tyr	cgc Arg	tcc Ser	ttc Phe 170	tgc Cys	aag Lys	atc Ile	ctg Leu	ttt Phe 175	cag Gln	gac Asp	tat Tyr	cac His	agt Ser 180	ttc Phe	703
gtg Val	act Thr	cat His	ggc Gly 185	tgc Cys	act Thr	gtg Val	gac Asp	aac Asn 190	ccc Pro	gtc Val	ctg Leu	gag Glu	cgg Arg 195	ttc Phe	atc Ile	751
tcc Ser	ctc Leu	ttc Phe 200	aac Asn	agc Ser	gtc Val	tca Ser	cag Gln 205	tgg Trp	gtg Val	cag Gln	ctc Leu	atg Met 210	atc Ile	ctc Leu	agc Ser	799
aaa Lys	ccc Pro 215	aca Thr	gcc Ala	ccg Pro	cag Gln	cgg Arg 220	gcc Ala	ctg Leu	gtc Val	atc Ile	aca Thr 225	cac His	ttt Phe	gtc Val	cac His	847
gtg Val 230	gcg Ala	gag Glu	aag Lys	ctg Leu	cta Leu 235	cag Gln	ctg Leu	cag Gln	aac Asn	ttc Phe 240	aac Asn	acg Thr	ctg Leu	atg Met	gca Ala 245	895
gtg Val	gtc Val	ggg Gly	ggc Gly	ctg Leu 250	agc Ser	cac His	agc Ser	tcc Ser	atc Ile 255	tcc Ser	cgc Arg	ctc Leu	aag Lys	gag Glu 260	acc Thr	943
cac His	agc Ser	cac His	gtt Val 265	agc Ser	cct Pro	gag Glu	acc Thr	atc Ile 270	aag Lys	ctc Leu	tgg Trp	gag Glu	ggt Gly 275	ctc Leu	acg Thr	991
gaa	cta	gtg	acg	gcg	aca	ggc	aac	tat	ggc	aac	tac	cgg	cgt	cgg	ctg	1039

Glu	Leu	Val	Thr	Ala	Thr	Gly	Asn	Tyr	Gly	Asn	Tyr	Arg	Arg	Arg	Leu	
		280					285					290				
gca	gcc	tgt	gtg	ggc	ttc	cgc	ttc	cgc	atc	ctg	ggg	gtg	cac	ctc	aag	1087
Ala	Ala	Cys	Val	Gly	Phe	Arg	Phe	Pro	Ile	Leu	Gly	Val	His	Leu	Lys	
	295					300					305					
gac	ctg	gtg	gcc	ctg	cag	ctg	gca	ctg	cct	gac	tgg	ctg	gac	cca	gcc	1135
Asp	Leu	Val	Ala	Leu	Gln	Leu	Ala	Leu	Pro	Asp	Trp	Leu	Asp	Pro	Ala	
310					315					320					325	
cgg	acc	cgg	ctc	aac	ggg	gcc	aag	atg	aag	cag	ctc	ttt	agc	atc	ctg	1183
Arg	Thr	Arg	Leu	Asn	Gly	Ala	Lys	Met	Lys	Gln	Leu	Phe	Ser	Ile	Leu	
				330					335					340		
gag	gag	ctg	gcc	atg	gtg	acc	agc	ctg	cgg	cca	cca	gta	cag	gcc	aac	1231
Glu	Glu	Leu	Ala	Met	Val	Thr	Ser	Leu	Arg	Pro	Pro	Val	Gln	Ala	Asn	
			345					350					355			
ccc	gac	ctg	ctg	agc	ctg	ctc	acg	gtg	tct	ctg	gat	cag	tat	cag	acg	1279
Pro	Asp	Leu	Leu	Ser	Leu	Leu	Thr	Val	Ser	Leu	Asp	Gln	Tyr	Gln	Thr	
		360					365					370				
gag	gat	gag	ctg	tac	cag	ctg	tcc	ctg	cag	cgg	gag	ccg	cgc	tcc	aag	1327
Glu	Asp	Glu	Leu	Tyr	Gln	Leu	Ser	Leu	Gln	Arg	Glu	Pro	Arg	Ser	Lys	
	375					380					385					
tct	tcg	cca	acc	agc	ccc	acg	agt	tgc	acc	cca	cca	ccc	cgg	ccc	ccg	1375
Ser	Ser	Pro	Thr	Ser	Pro	Thr	Ser	Cys	Thr	Pro	Pro	Pro	Arg	Pro	Pro	
390					395				400					405		
gta	ttg	gag	gag	tgg	acc	tcg	gct	gcc	aaa	ccc	aag	ctg	gat	cag	gcc	1423
Val	Leu	Glu	Glu	Trp	Thr	Ser	Ala	Ala	Lys	Pro	Lys	Leu	Asp	Gln	Ala	
				410				415						420		
ctc	gtg	gtg	gag	cac	atc	gag	aag	atg	gtg	gag	tct	gtg	ttc	cgg	aac	1471
Leu	Val	Val	Glu	His	Ile	Glu	Lys	Met	Val	Glu	Ser	Val	Phe	Arg	Asn	
			425					430					435			
ttt	gac	gtc	gat	ggg	gat	ggc	cac	atc	tca	cag	gaa	gaa	ttc	cag	atc	1519
Phe	Asp	Val	Asp	Gly	Asp	Gly	His	Ile	Ser	Gln	Glu	Glu	Phe	Gln	Ile	
		440				445						450				
atc	cgt	ggg	aac	ttc	cct	tac	ctc	agc	gcc	ttt	ggg	gac	ctc	gac	cag	1567
Ile	Arg	Gly	Asn	Phe	Pro	Tyr	Leu	Ser	Ala	Phe	Gly	Asp	Leu	Asp	Gln	
	455					460					465					
aac	cag	gat	ggc	tgc	atc	agc	agg	gag	gag	atg	gtt	tcc	tat	ttc	ctg	1615
Asn	Gln	Asp	Gly	Cys	Ile	Ser	Arg	Glu	Glu	Met	Val	Ser	Tyr	Phe	Leu	
470					475					480					485	
cgc	tcc	agc	tct	gtg	ttg	ggg	ggg	cgc	atg	ggc	ttc	gta	cac	aac	ttc	1663
Arg	Ser	Ser	Ser	Val	Leu	Gly	Gly	Arg	Met	Gly	Phe	Val	His	Asn	Phe	
				490					495					500		
cag	gag	agc	aac	tcc	ttg	cgc	ccc	gtc	gcc	tgc	cgc	cac	tgc	aaa	gcc	1711
Gln	Glu	Ser	Asn	Ser	Leu	Arg	Pro	Val	Ala	Cys	Arg	His	Cys	Lys	Ala	
			505					510					515			
ctg	atc	ctg	ggc	atc	tac	aag	cag	ggc	ctc	aaa	tgc	cga	gcc	tgt	gga	1759
Leu	Ile	Leu	Gly	Ile	Tyr	Lys	Gln	Gly	Leu	Lys	Cys	Arg	Ala	Cys	Gly	
		520					525					530				
gtg	aac	tgc	cac	aag	cag	tgc	aag	gat	cgc	ctg	tca	gtt	gag	tgt	cgg	1807
Val	Asn	Cys	His	Lys	Gln	Cys	Lys	Asp	Arg	Leu	Ser	Val	Glu	Cys	Arg	
	535					540					545					
cgc	agg	gcc	cag	agt	gtg	agc	ctg	gag	ggg	tct	gca	ccc	tca	ccc	tca	1855
Arg	Arg	Ala	Gln	Ser	Val	Ser	Leu	Glu	Gly	Ser	Ala	Pro	Ser	Pro	Ser	

```

<210> 4
<211> 609
<212> PRT
<213> Homo sapiens

<400> 4
Met Ala Gly Thr Leu5 Asp Leu Asp Lys Gly10 Cys Thr Val Glu Glu15 Leu
1
Leu Arg Gly Cys20 Ile Glu Ala Phe Asp25 Asp Ser Gly Lys Val30 Arg Asp
Pro Gln Leu35 Val Arg Met Phe Leu40 Met Met His Pro Trp45 Tyr Ile Pro
Ser Ser50 Gln Leu Ala Ala Lys55 Leu Leu His Ile Tyr60 Gln Gln Ser Arg
Lys65 Asp Asn Ser Asn Ser70 Leu Gln Val Lys Thr75 Cys His Leu Val Arg80
Tyr Trp Ile Ser Ala85 Phe Pro Ala Glu Phe90 Asp Leu Asn Pro Glu95 Leu
Ala Glu Gln Ile100 Lys Glu Leu Lys Ala105 Leu Leu Asp Gln Glu110 Gly Asn
Arg Arg His115 Ser Ser Leu Ile Asp120 Ile Asp Ser Val Pro125 Thr Tyr Lys
Trp Lys130 Arg Gln Val Thr Gln135 Arg Asn Pro Val Gly140 Gln Lys Lys Arg
Lys145 Met Ser Leu Leu Phe150 Asp His Leu Glu Pro155 Met Glu Leu Ala Glu160
His Leu Thr Tyr Leu165 Glu Tyr Arg Ser Phe170 Cys Lys Ile Leu Phe175 Gln
Asp Tyr His Ser180 Phe Val Thr His Gly185 Cys Thr Val Asp Asn190 Pro Val
Leu Glu Arg195 Phe Ile Ser Leu Phe200 Asn Ser Val Ser Gln205 Trp Val Gln
Leu Met Ile Leu Ser Lys Pro Thr Ala Pro Gln Arg Ala Leu Val Ile

```

210					215					220				
Thr 225	His	Phe	Val	His 230	Val	Ala	Glu	Lys	Leu 235	Gln	Leu	Gln	Asn	Phe 240
Asn	Thr	Leu	Met	Ala 245	Val	Val	Gly	Gly	Leu 250	Ser	His	Ser	Ser	Ile 255
Arg	Leu	Lys	Glu 260	Thr	His	Ser	His	Val 265	Ser	Pro	Glu	Thr	Ile 270	Lys 275
Trp	Glu	Gly 275	Leu	Thr	Glu	Leu	Val 280	Thr	Ala	Thr	Gly	Asn 285	Tyr	Gly 290
Tyr	Arg 290	Arg	Arg	Leu	Ala	Ala 295	Cys	Val	Gly	Phe	Arg 300	Phe	Pro	Ile 305
Gly 305	Val	His	Leu	Lys	Asp 310	Leu	Val	Ala	Leu 315	Gln	Leu	Ala	Leu	Pro 320
Trp	Leu	Asp	Pro	Ala 325	Arg	Thr	Arg	Leu	Asn 330	Gly	Ala	Lys	Met	Lys 335
Leu	Phe	Ser	Ile 340	Leu	Glu	Glu	Leu	Ala 345	Met	Val	Thr	Ser	Leu	Arg 350
Pro	Val	Gln 355	Ala	Asn	Pro	Asp	Leu 360	Leu	Ser	Leu	Leu	Thr 365	Val	Ser 370
Asp	Gln 370	Tyr	Gln	Thr	Glu	Asp 375	Glu	Leu	Tyr	Gln 380	Leu	Ser	Leu	Gln 385
Glu 385	Pro	Arg	Ser	Lys	Ser 390	Ser	Pro	Thr	Ser	Pro 395	Thr	Ser	Cys	Thr 400
Pro	Pro	Arg	Pro	Pro 405	Val	Leu	Glu	Glu	Trp 410	Thr	Ser	Ala	Ala	Lys 415
Lys	Leu	Asp	Gln 420	Ala	Leu	Val	Val 425	Glu	His	Ile	Glu	Lys	Met 430	Val 435
Ser	Val	Phe 435	Arg	Asn	Phe	Asp	Val 440	Asp	Gly	Asp	Gly	His 445	Ile	Ser 450
Glu	Glu 450	Phe	Gln	Ile	Ile	Arg 455	Gly	Asn	Phe	Pro	Tyr 460	Leu	Ser	Ala 465
Gly 465	Asp	Leu	Asp	Gln	Asn 470	Gln	Asp	Gly	Cys	Ile 475	Ser	Arg	Glu	Met 480
Val	Ser	Tyr	Phe	Leu 485	Arg	Ser	Ser	Ser	Val 490	Leu	Gly	Gly	Arg	Met 495
Phe	Val	His	Asn 500	Phe	Gln	Glu	Ser	Asn 505	Ser	Leu	Arg	Pro	Val 510	Ala 515
Arg	His	Cys 515	Lys	Ala	Leu	Ile	Leu 520	Gly	Ile	Tyr	Lys	Gln 525	Gly	Leu 530
Cys	Arg 530	Ala	Cys	Gly	Val	Asn 535	Cys	His	Lys	Gln	Cys 540	Lys	Asp	Arg 545
Ser 545	Val	Glu	Cys	Arg	Arg 550	Arg	Ala	Gln	Ser	Val 555	Ser	Leu	Glu	Gly 560
Ala	Pro	Ser	Pro	Ser 565	Pro	Met	His	Ser	His 570	His	His	Arg	Ala	Phe 575
Phe	Ser	Leu	Pro	Arg	Pro	Gly	Arg	Arg	Gly	Ser	Arg	Pro	Pro	Glu 580

```

      580              585              590
Arg Glu Glu Glu Val Gln Thr Val Glu Asp Gly Val Phe Asp Ile His
    595                    600                    605

Leu

<210> 5
<211> 3624
<212> DNA
<213> Rattus norvegicus

<220>
<221> CDS
<222> (145)..(2529)
<223> CalDAG-GEFII

<400> 5
gcgcctgggt cggtctgccg ggctccgaga gtggccggct ccgggctgcg agggccggcg 60
gcggccgagg gatgcgccgc tcccggcga gccgatcca ccgctccagg tgagacggct 120
ccaggggcmc agagagccgc ggcc atg gga acc ctg ggc aag gcg aga gag 171
                        Met Gly Thr Leu Gly Lys Ala Arg Glu
                          1                      5

gct ccg cgg aaa cct tgc cat ggc tcc aga gct ggc ccc aaa gga aga 219
Ala Pro Arg Lys Pro Cys His Gly Ser Arg Ala Gly Pro Lys Gly Arg
   10                15               20                25

cta gag gcc aaa tca acc aac agt cct ctc cct gcc cag ccc agc ttg 267
Leu Glu Ala Lys Ser Thr Asn Ser Pro Leu Pro Ala Gln Pro Ser Leu
                     30                 35                  40

gcc cag atc acc cag ttc cga atg atg gtg tcc ctg gga cat ctg gcc 315
Ala Gln Ile Thr Gln Phe Arg Met Met Val Ser Leu Gly His Leu Ala
                   45                50                55

aaa gga gcc agc ctg gat gat ctt att gac agc tgc att caa tct ttc 363
Lys Gly Ala Ser Leu Asp Asp Leu Ile Asp Ser Cys Ile Gln Ser Phe
          60                                70

gat gca gat gga aac ctg tgt cga agt aac cag ctg tta caa gtc atg 411
Asp Ala Asp Gly Asn Leu Cys Arg Ser Asn Gln Leu Leu Gln Val Met
       75                80                85

cta acc atg cac cga atc atc atc tcc tcg gcc gag ctg ctg caa aaa 459
Leu Thr Met His Arg Ile Ile Ile Ser Ser Ala Glu Leu Leu Gln Lys
   90                95                100             105

ctc atg aat cta tat aag gac gcc ctg gaa aag aat tct cca gga att 507
Leu Met Asn Leu Tyr Lys Asp Ala Leu Glu Lys Asn Ser Pro Gly Ile
                     110                115             120

tgc ttg aag atc tgc tat ttt gtc agg tat tgg ata aca gaa ttc tgg 555
Cys Leu Lys Ile Cys Tyr Phe Val Arg Tyr Trp Ile Thr Glu Phe Trp
        125                130              135

atc atg ttc aaa atg gat gcc agc ttg acc agc acc atg gaa gag ttt 603
Ile Met Phe Lys Met Asp Ala Ser Leu Thr Ser Thr Met Glu Glu Phe
          140                145             150

cag gac ctg gtg aaa gcc aat ggt gag gag tcc cac tgc cac ctc atc 651
Gln Asp Leu Val Lys Ala Asn Gly Glu Glu Ser His Cys His Leu Ile
      155                160              165

gac acg aca caa att aat tct cga gac tgg tcc agg aaa ctg act cag 699

```

Asp 170	Thr	Thr	Gln	Ile	Asn 175	Ser	Arg	Asp	Trp	Ser 180	Arg	Lys	Leu	Thr	Gln 185	
agg Arg	ata Ile	aaa Lys	tca Ser	aat Asn 190	acc Thr	agc Ser	aag Lys	aag Lys	cgg Arg 195	aaa Lys	gtg Val	tcc Ser	ctg Leu	ctg Leu 200	ttt Phe	747
gac Asp	cat His	ctt Leu	gaa Glu 205	cct Pro	gaa Glu	gaa Glu	ctg Leu	tct Ser 210	gaa Glu	cac His	ctc Leu	acc Thr	tac Tyr 215	ctt Leu	gag Glu	795
ttc Phe	aag Lys	tcc Ser 220	ttc Phe	cga Arg	cgg Arg	ata Ile	tct Ser 225	ttc Phe	tct Ser	gat Asp	tat Tyr	cag Gln 230	aat Asn	tac Tyr	ctt Leu	843
gta Val	aac Asn 235	agc Ser	tgc Cys	gta Val	aag Lys	gag Glu 240	aac Asn	ccc Pro	acc Thr	atg Met	gag Glu 245	cgg Arg	tcc Ser	att Ile	gcc Ala	891
ctg Leu 250	tgc Cys	aat Asn	ggc Gly	atc Ile	tcc Ser 255	cag Gln	tgg Trp	gta Val	caa Gln	ctg Leu 260	atg Met	gtt Val	ctc Leu	agc Ser	cgt Arg 265	939
ccc Pro	acc Thr	cca Pro	cag Gln	ctc Leu 270	cgg Arg	gca Ala	gag Glu	gtc Val	ttc Phe 275	atc Ile	aag Lys	ttc Phe	atc Ile	cat His 280	gtg Val	987
gct Ala	cag Gln	aag Lys	ctc Leu 285	cac His	cag Gln	cta Leu	cag Gln	aac Asn 290	ttc Phe	aac Asn	acg Thr	cta Leu	atg Met 295	gct Ala	gtg Val	1035
atc Ile	ggg Gly	gga Gly 300	ctg Leu	tgt Cys	cac His	agc Ser	tcc Ser 305	atc Ile	tcc Ser	agg Arg	ctc Leu	aag Lys 310	gag Glu	aca Thr	agt Ser	1083
tca Ser	cat His 315	gtc Val	cca Pro	cat His	gag Glu	atc Ile 320	aat Asn	aag Lys	gtt Val	ctg Leu	ggg Gly 325	gag Glu	atg Met	act Thr	gaa Glu	1131
ctg Leu 330	ctg Leu	tcc Ser	tcc Ser	tgc Cys	aga Arg 335	aac Asn	tat Tyr	gac Asp	aac Asn	tac Tyr 340	agg Arg	cga Arg	gcc Ala	tat Tyr	ggg Gly 345	1179
gag Glu	tgc Cys	acc Thr	cac His	ttc Phe 350	aaa Lys	atc Ile	ccc Pro	ata Ile	ctg Leu 355	ggg Gly	gtg Val	cac His	ctc Leu	aag Lys 360	gac Asp	1227
ctc Leu	ata Ile	tcc Ser	cta Leu 365	tat Tyr	gaa Glu	gcc Ala	atg Met	cct Pro 370	gac Asp	tac Tyr	ctg Leu	gaa Glu	gac Asp 375	ggg Gly	aag Lys	1275
gtg Val	aat Asn	gtc Val 380	caa Gln	aag Lys	ctc Leu	ctg Leu	gcc Ala 385	ctt Leu	tac Tyr	aat Asn	cac His	atc Ile 390	aat Asn	gag Glu	ttg Leu	1323
gtc Val	cag Gln 395	ctg Leu	cag Gln	gac Asp	gtg Val	gcc Ala 400	cca Pro	cca Pro	ttg Leu	gat Asp	gcc Ala 405	aac Asn	aag Lys	gac Asp	ctg Leu	1371
gtg Val 410	cac His	ctg Leu	ctg Leu	acg Thr	tta Leu 415	tcc Ser	ctg Leu	gat Asp	cta Leu	tac Tyr 420	tac Tyr	acc Thr	gaa Glu	gat Asp	gaa Glu 425	1419
atc Ile	tat Tyr	gag Glu	ctt Leu	tcc Ser 430	tac Tyr	gcc Ala	cgt Arg	gaa Glu	cca Pro 435	agg Arg	aac Asn	cac His	agg Arg	gcc Ala 440	ccg Pro	1467
cca Pro	ctg Leu	aca Thr	cct Pro	tcg Ser	aag Lys	cca Pro	cca Pro	gtt Val	gta Val	gtg Val	gac Asp	tgg Trp	gcc Ala	tct Ser	gga Gly	1515

			445				450				455							
gtg Val	tct Ser	ccc Pro 460	aaa Lys	cct Pro	gac Asp	cgc Pro	aag Lys 465	acc Thr	atc Ile	agc Ser	aaa Lys	cac His 470	gtc Val	caa Gln	agg Arg	1563		
atg Met	gtg Val 475	gat Asp	tct Ser	gtc Val	ttt Phe	aag Lys 480	aac Asn	tat Tyr	gat Asp	ctc Leu	gac Asp 485	cag Gln	gat Asp	gga Gly	tat Tyr	1611		
atc Ile 490	tct Ser	cag Gln	gag Glu	gag Glu	ttt Phe 495	gaa Glu	aag Lys	att Ile	gct Ala	gcg Ala 500	agc Ser	ttt Phe	cca Pro	ttt Phe	tcc Ser 505	1659		
ttc Phe	tgt Cys	gtg Val	atg Met	gac Asp 510	aaa Lys	gat Asp	agg Arg	gag Glu	ggc Gly 515	ctc Leu	atc Ile	agc Ser	aga Arg	gac Asp 520	gag Glu	1707		
atc Ile	aca Thr	gcc Ala	tac Tyr 525	ttc Phe	atg Met	agg Arg	gcc Ala	agc Ser 530	tcc Ser	atc Ile	tat Tyr	tcc Ser	aag Lys 535	ctg Leu	ggc Gly	1755		
cta Leu	ggc Gly	ttt Phe 540	cca Pro	cac His	aac Asn	ttt Phe	caa Gln 545	gag Glu	acc Thr	act Thr	tac Tyr	ctg Leu 550	aag Lys	ccc Pro	acc Thr	1803		
ttc Phe	tgt Cys 555	gac Asp	aac Asn	tgt Cys	gct Ala	ggc Gly 560	ttt Phe	ctc Leu	tgg Trp	ggt Gly	gtg Val 565	atc Ile	aag Lys	caa Gln	ggc Gly	1851		
tat Tyr 570	cgc Arg	tgt Cys	aaa Lys	gac Asp	tgt Cys 575	ggg Gly	atg Met	aac Asn	tgc Cys	cac His 580	aaa Lys	cag Gln	tgc Cys	aaa Lys	gac Asp 585	1899		
ctg Leu	gta Val	gtg Val	ttt Phe	gag Glu 590	tgc Cys	aag Lys	aaa Lys	cga Arg	tcc Ser 595	aag Lys	agc Ser	ccg Pro	gcg Ala	gta Val 600	tcc Ser	1947		
aca Thr	gaa Glu	aac Asn	atc Ile 605	agc Ser	tct Ser	gtg Val	gtg Val	cca Pro 610	atg Met	tcc Ser	act Thr	ctt Leu	tgt Cys 615	cca Pro	ctg Leu	1995		
gga Gly	acc Thr	aaa Lys 620	gat Asp	ctg Leu	ctc Leu	cat His	gca Ala 625	ccc Pro	gaa Glu	gaa Glu	gga Gly	tct Ser 630	ttc Phe	att Ile	ttc Phe	2043		
cag Gln	aat Asn 635	gga Gly	gag Glu	gtt Val	gtg Val	gac Asp 640	cac His	agt Ser	gag Glu	gag Glu	agc Ser 645	aag Lys	gat Asp	agg Arg	acc Thr	2091		
atc Ile 650	atg Met	ctc Leu	ttg Leu	ggc Gly	gta Val 655	tcc Ser	tca Ser	cag Gln	aaa Lys	att Ile 660	tca Ser	gtt Val	cgg Arg	ctg Leu	aag Lys 665	2139		
agg Arg	act Thr	gtt Val	gcc Ala	cac His 670	aag Lys	acc Thr	acc Thr	cag Gln	aca Thr 675	gaa Glu	tca Ser	ttt Phe	cct Pro	tgg Trp 680	gtt Val	2187		
ggc Gly	ggc Gly	gag Glu 685	atg Met	ccc Pro	cct Pro	ggt Gly	cac His	ttt Phe 690	gtg Val	ctg Leu	act Thr	tct Ser	cca Pro 695	aga Arg	aag Lys	2235		
tca Ser	gca Ala	caa Gln 700	ggg Gly	gct Ala	ctt Leu	tat Tyr	gtg Val 705	cac His	agt Ser	cca Pro	gca Ala	tct Ser 710	ccg Pro	tgc Cys	ccc Pro	2283		
agc Ser	cca Pro 715	gca Ala	ctg Leu	gtc Val	cgg Arg	aag Lys 720	cgg Arg	gca Ala	ttc Phe	gtc Val	aag Lys 725	tgg Trp	gag Glu	aac Asn	aaa Lys	2331		

gag tcc ctt atc aaa cca aaa cca gag ctt cac ctc agg ctc cgg acc 2379
Glu Ser Leu Ile Lys Pro Lys Pro Glu Leu His Leu Arg Leu Arg Thr 745
730 735 740

tac caa gaa ctg gaa cag gag gta aat acc ctg agg gca gat aac gat 2427
Tyr Gln Glu Leu Glu Gln Glu Val Asn Thr Leu Arg Ala Asp Asn Asp 760
750 755

gct ctg aag atc cag ctg aag tat gca cag aaa caa ata gaa tcc ctg 2475
Ala Leu Lys Ile Gln Leu Lys Tyr Ala Gln Lys Gln Ile Glu Ser Leu 775
765 770

cag ctt ggc aaa agc aat cac gtc tta gca cag atg gac cac ggt gat 2523
Gln Leu Gly Lys Ser Asn His Val Leu Ala Gln Met Asp His Gly Asp 790
780 785 790

ggt act taatccagaa attcaaggaa cagaatctgc agacgggttt actgggatct 2579
Gly Thr 795

cacttcaaaa ctgattgcag aggttcagca acttttagact gattgacttt taaagggcag 2639
agatagccac tggtattggt gtccttggtt ttttccctaa ctgtctttta tgtgagtcgt 2699
gggtttttcag tcagttgagt aaaagggaag aaaagttcca gcatgtaaaa cactgagcag 2759
tcatattcta accttttctt tcttttactg aaaccaatac caaacatgtg ctaaaataag 2819
agtatagttt cctgacagtg tctcaggcag cctgctttac tgattcgcac ttagataaag 2879
agccccctggt gaatcaattt gctgccttcc ccagggtttc tgcaaaactgg gagttacttt 2939
gttctacccg aaaacctgct cataatggaa acaataccct aaaagtgggtg actttgacat 2999
gcctgcattg tttgtgaagc tgatccctac cttattcact acctcagatc tcaagagcct 3059
ctttccctgt cctttacttg cccctatatt cctcccttgc ttgtaggcac atcagcatat 3119
ctacctagaa gtgacctcct agagatgtag cctgtgttta aatccagagc ttcttaattg 3179
aacttgacat tgtctgattg caggccaact tatagctcta ggtcttagtc attatgaaat 3239
acaaaaataaa actaccata atcatcaatt accatgtcat acacagaact cattatctaa 3299
gtcaaacgac gtaaacacgc ctgtccagat agtctttctt tagtggatag aattaacaat 3359
ctgtgacctt ttaaaagaca tgggtctgtga acagtgttta gttcttcaag ttctagtctc 3419
tctctctctc tctctctctc tctctctctc tctctctctc tctctcaagt aaactatcgt 3479
aaggtcctct tttgaactga atctgtgctt aaaattgtct tcaacttttat tatctacaaa 3539
taagctatgg gaggcgatgg ctttaacagc tgacagcatt tacctatgtg tagaatatgt 3599
gtatatagta tcaggcatgc atgga 3624

<210> 6
<211> 795
<212> PRT
<213> Rattus norvegicus

<400> 6
Met Gly Thr Leu Gly Lys Ala Arg Glu Ala Pro Arg Lys Pro Cys His
1 5 10 15
Gly Ser Arg Ala Gly Pro Lys Gly Arg Leu Glu Ala Lys Ser Thr Asn
20 25 30

Ser	Pro	Leu 35	Pro	Ala	Gln	Pro	Ser 40	Leu	Ala	Gln	Ile	Thr 45	Gln	Phe	Arg
Met	Met 50	Val	Ser	Leu	Gly	His 55	Leu	Ala	Lys	Gly	Ala 60	Ser	Leu	Asp	Asp
Leu 65	Ile	Asp	Ser	Cys	Ile 70	Gln	Ser	Phe	Asp	Ala 75	Asp	Gly	Asn	Leu	Cys 80
Arg	Ser	Asn	Gln	Leu 85	Leu	Gln	Val	Met	Leu 90	Thr	Met	His	Arg	Ile 95	Ile
Ile	Ser	Ser	Ala 100	Glu	Leu	Leu	Gln	Lys 105	Leu	Met	Asn	Leu	Tyr 110	Lys	Asp
Ala	Leu	Glu 115	Lys	Asn	Ser	Pro	Gly 120	Ile	Cys	Leu	Lys	Ile 125	Cys	Tyr	Phe
Val	Arg 130	Tyr	Trp	Ile	Thr	Glu 135	Phe	Trp	Ile	Met	Phe 140	Lys	Met	Asp	Ala
Ser 145	Leu	Thr	Ser	Thr	Met 150	Glu	Glu	Phe	Gln	Asp 155	Leu	Val	Lys	Ala	Asn 160
Gly	Glu	Glu	Ser	His 165	Cys	His	Leu	Ile	Asp 170	Thr	Thr	Gln	Ile	Asn 175	Ser
Arg	Asp	Trp	Ser 180	Arg	Lys	Leu	Thr	Gln 185	Arg	Ile	Lys	Ser	Asn 190	Thr	Ser
Lys	Lys	Arg 195	Lys	Val	Ser	Leu	Leu 200	Phe	Asp	His	Leu	Glu 205	Pro	Glu	Glu
Leu	Ser 210	Glu	His	Leu	Thr	Tyr 215	Leu	Glu	Phe	Lys	Ser 220	Phe	Arg	Arg	Ile
Ser 225	Phe	Ser	Asp	Tyr	Gln 230	Asn	Tyr	Leu	Val	Asn 235	Ser	Cys	Val	Lys	Glu 240
Asn	Pro	Thr	Met	Glu 245	Arg	Ser	Ile	Ala	Leu 250	Cys	Asn	Gly	Ile	Ser 255	Gln
Trp	Val	Gln	Leu 260	Met	Val	Leu	Ser	Arg 265	Pro	Thr	Pro	Gln	Leu 270	Arg	Ala
Glu	Val	Phe 275	Ile	Lys	Phe	Ile	His 280	Val	Ala	Gln	Lys	Leu 285	His	Gln	Leu
Gln	Asn 290	Phe	Asn	Thr	Leu	Met 295	Ala	Val	Ile	Gly	Gly 300	Leu	Cys	His	Ser
Ser 305	Ile	Ser	Arg	Leu	Lys 310	Glu	Thr	Ser	Ser	His 315	Val	Pro	His	Glu	Ile 320
Asn	Lys	Val	Leu	Gly 325	Glu	Met	Thr	Glu	Leu 330	Leu	Ser	Ser	Cys	Arg 335	Asn
Tyr	Asp	Asn	Tyr 340	Arg	Arg	Ala	Tyr	Gly 345	Glu	Cys	Thr	His	Phe 350	Lys	Ile
Pro	Ile	Leu 355	Gly	Val	His	Leu	Lys 360	Asp	Leu	Ile	Ser	Leu 365	Tyr	Glu	Ala
Met	Pro 370	Asp	Tyr	Leu	Glu	Asp 375	Gly	Lys	Val	Asn	Val 380	Gln	Lys	Leu	Leu
Ala 385	Leu	Tyr	Asn	His	Ile 390	Asn	Glu	Leu	Val	Gln 395	Leu	Gln	Asp	Val	Ala 400

Pro	Pro	Leu	Asp	Ala 405	Asn	Lys	Asp	Leu	Val 410	His	Leu	Leu	Thr	Leu 415	Ser
Leu	Asp	Leu	Tyr 420	Tyr	Thr	Glu	Asp	Glu 425	Ile	Tyr	Glu	Leu	Ser 430	Tyr	Ala
Arg	Glu	Pro 435	Arg	Asn	His	Arg	Ala 440	Pro	Pro	Leu	Thr	Pro 445	Ser	Lys	Pro
Pro	Val 450	Val	Val	Asp	Trp	Ala 455	Ser	Gly	Val	Ser	Pro 460	Lys	Pro	Asp	Pro
Lys 465	Thr	Ile	Ser	Lys	His 470	Val	Gln	Arg	Met	Val 475	Asp	Ser	Val	Phe	Lys 480
Asn	Tyr	Asp	Leu	Asp 485	Gln	Asp	Gly	Tyr	Ile 490	Ser	Gln	Glu	Glu	Phe 495	Glu
Lys	Ile	Ala	Ala 500	Ser	Phe	Pro	Phe	Ser 505	Phe	Cys	Val	Met	Asp 510	Lys	Asp
Arg	Glu	Gly 515	Leu	Ile	Ser	Arg	Asp 520	Glu	Ile	Thr	Ala	Tyr 525	Phe	Met	Arg
Ala	Ser 530	Ser	Ile	Tyr	Ser	Lys 535	Leu	Gly	Leu	Gly	Phe 540	Pro	His	Asn	Phe
Gln 545	Glu	Thr	Thr	Tyr	Leu 550	Lys	Pro	Thr	Phe	Cys 555	Asp	Asn	Cys	Ala	Gly 560
Phe	Leu	Trp	Gly	Val 565	Ile	Lys	Gln	Gly	Tyr 570	Arg	Cys	Lys	Asp	Cys 575	Gly
Met	Asn	Cys	His 580	Lys	Gln	Cys	Lys	Asp 585	Leu	Val	Val	Phe	Glu 590	Cys	Lys
Lys	Arg	Ser 595	Lys	Ser	Pro	Ala	Val 600	Ser	Thr	Glu	Asn 605	Ile	Ser	Ser	Val
Val 610	Pro	Met	Ser	Thr	Leu	Cys 615	Pro	Leu	Gly	Thr	Lys 620	Asp	Leu	Leu	His
Ala 625	Pro	Glu	Glu	Gly	Ser 630	Phe	Ile	Phe	Gln	Asn 635	Gly	Glu	Val	Val	Asp 640
His	Ser	Glu	Glu	Ser 645	Lys	Asp	Arg	Thr	Ile 650	Met	Leu	Leu	Gly	Val 655	Ser
Ser	Gln	Lys	Ile 660	Ser	Val	Arg	Leu	Lys 665	Arg	Thr	Val	Ala	His 670	Lys	Thr
Thr	Gln	Thr 675	Glu	Ser	Phe	Pro	Trp 680	Val	Gly	Gly	Glu	Met 685	Pro	Pro	Gly
His 690	Phe	Val	Leu	Thr	Ser	Pro 695	Arg	Lys	Ser	Ala	Gln 700	Gly	Ala	Leu	Tyr
Val 705	His	Ser	Pro	Ala	Ser 710	Pro	Cys	Pro	Ser	Pro 715	Ala	Leu	Val	Arg	Lys 720
Arg	Ala	Phe	Val	Lys 725	Trp	Glu	Asn	Lys	Glu 730	Ser	Leu	Ile	Lys	Pro 735	Lys
Pro	Glu	Leu	His 740	Leu	Arg	Leu	Arg	Thr 745	Tyr	Gln	Glu	Leu	Glu 750	Gln	Glu
Val	Asn	Thr 755	Leu	Arg	Ala	Asp	Asn	Asp 760	Ala	Leu	Lys	Ile 765	Gln	Leu	Lys

1000
 900
 800
 700
 600
 500
 400
 300
 200
 100
 0

Tyr Ala Gln Lys Gln Ile Glu Ser Leu Gln Leu Gly Lys Ser Asn His
770 775 780

Val Leu Ala Gln Met Asp His Gly Asp Gly Thr
785 790 795

<210> 7
<211> 5075
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (104)..(2494)
<223> CalDAG-GEFII

<400> 7
cccggggggg agagagccgg caggcggcgg cggtggtggc gggggcgatg cgccgcgccc 60
ggccgcgcta ggtgagccgg caccgggagc gcgggcccgcg gcc atg ggc acc ctg 115
Met Gly Thr Leu
1
ggc aag gcg aga gag gct ccg cgg aaa cct tcc cat ggc tgc aga gct 163
Gly Lys Ala Arg Glu Ala Pro Arg Lys Pro Ser His Gly Cys Arg Ala
5 10 15 20
gcc tct aaa gca aga cta gag gca aag cca gcc aac agc ccc ttc ccc 211
Ala Ser Lys Ala Arg Leu Glu Ala Lys Pro Ala Asn Ser Pro Phe Pro
25 30 35
tcc cat ccc agc ttg gcc cac atc acc cag ttc cga atg atg gtg tct 259
Ser His Pro Ser Leu Ala His Ile Thr Gln Phe Arg Met Met Val Ser
40 45 50
ctg gga cat tta gcc aaa gga gcc agc ctg gac gat ctc att gac agc 307
Leu Gly His Leu Ala Lys Gly Ala Ser Leu Asp Asp Leu Ile Asp Ser
55 60 65
tgc att caa tct ttt gat gca gat gga aac ctg tgt cga agt aac caa 355
Cys Ile Gln Ser Phe Asp Ala Asp Gly Asn Leu Cys Arg Ser Asn Gln
70 75 80
ctg ttg caa gtc atg ctg acc atg cac cga att gtc atc tcc tct gca 403
Leu Leu Gln Val Met Leu Thr Met His Arg Ile Val Ile Ser Ser Ala
85 90 95 100
gaa ctg ctc caa aaa gtt atc acc ctc tat aag gat gct ttg gca aag 451
Glu Leu Leu Gln Lys Val Ile Thr Leu Tyr Lys Asp Ala Leu Ala Lys
105 110 115
aat tca cca gga ctt tgc ctg aag atc tgt tat ttt gta agg tat tgg 499
Asn Ser Pro Gly Leu Cys Leu Lys Ile Cys Tyr Phe Val Arg Tyr Trp
120 125 130
ata aca gaa ttc tgg gtc atg ttt aaa atg gac gcc agc ttg aca gac 547
Ile Thr Glu Phe Trp Val Met Phe Lys Met Asp Ala Ser Leu Thr Asp
135 140 145
act atg gag gag ttt cag gaa ctg gtg aaa gct aag ggt gag gag tta 595
Thr Met Glu Glu Phe Gln Glu Leu Val Lys Ala Lys Gly Glu Glu Leu
150 155 160
cat tgc cgc ctg att gac aca act caa atc aat gcc cgt gac tgg tcc 643
His Cys Arg Leu Ile Asp Thr Thr Gln Ile Asn Ala Arg Asp Trp Ser
165 170 175 180

118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000

agg	aaa	ctt	act	caa	agg	ata	aaa	tca	aat	acc	agc	aag	aaa	cgg	aaa	691
Arg	Lys	Leu	Thr	Gln	Arg	Ile	Lys	Ser	Asn	Thr	Ser	Lys	Lys	Arg	Lys	
				185					190					195		
gtc	tcc	ctg	ctc	ttt	gac	cat	ctg	gaa	cca	gaa	gag	cta	tcc	gag	cac	739
Val	Ser	Leu	Leu	Phe	Asp	His	Leu	Glu	Pro	Glu	Glu	Leu	Ser	Glu	His	
			200					205					210			
ctc	acc	tac	ctt	gag	ttc	aag	tct	ttc	cgg	agg	ata	tcg	ttc	tct	gat	787
Leu	Thr	Tyr	Leu	Glu	Phe	Lys	Ser	Phe	Arg	Arg	Ile	Ser	Phe	Ser	Asp	
		215					220					225				
tat	cag	aat	tac	ctt	gta	aat	agc	tgt	gtg	aag	gaa	aac	ccc	acc	atg	835
Tyr	Gln	Asn	Tyr	Leu	Val	Asn	Ser	Cys	Val	Lys	Glu	Asn	Pro	Thr	Met	
	230					235					240					
gag	cga	tct	att	gct	ctg	tgc	aac	ggc	atc	tcc	cag	tgg	gta	caa	ctg	883
Glu	Arg	Ser	Ile	Ala	Leu	Cys	Asn	Gly	Ile	Ser	Gln	Trp	Val	Gln	Leu	
245					250					255					260	
atg	gtt	ctc	agc	cgc	ccc	acg	ccg	cag	ctc	cga	gca	gaa	gtc	ttc	atc	931
Met	Val	Leu	Ser	Arg	Pro	Thr	Pro	Gln	Leu	Arg	Ala	Glu	Val	Phe	Ile	
				265					270					275		
aag	ttc	atc	cag	gtg	gct	cag	aag	ctc	cac	caa	cta	cag	aac	ttc	aat	979
Lys	Phe	Ile	Gln	Val	Ala	Gln	Lys	Leu	His	Gln	Leu	Gln	Asn	Phe	Asn	
			280					285					290			
aca	ctg	atg	gct	gtg	ata	ggc	ggg	ctg	tgt	cac	agc	tca	atc	tcg	agg	1027
Thr	Leu	Met	Ala	Val	Ile	Gly	Gly	Leu	Cys	His	Ser	Ser	Ile	Ser	Arg	
		295					300					305				
ctc	aag	gag	aca	agt	tcg	cat	gtc	cca	cat	gaa	atc	aat	aag	gtt	ctc	1075
Leu	Lys	Glu	Thr	Ser	Ser	His	Val	Pro	His	Glu	Ile	Asn	Lys	Val	Leu	
	310					315					320					
ggc	gag	atg	act	gag	ctg	ctg	tcc	tcc	tcc	aga	aac	tac	gac	aat	tac	1123
Gly	Glu	Met	Thr	Glu	Leu	Leu	Ser	Ser	Ser	Arg	Asn	Tyr	Asp	Asn	Tyr	
325					330					335					340	
cgg	cga	gcc	tat	gga	gag	tgc	acc	gac	ttc	aag	atc	ccc	att	ctg	ggc	1171
Arg	Arg	Ala	Tyr	Gly	Glu	Cys	Thr	Asp	Phe	Lys	Ile	Pro	Ile	Leu	Gly	
				345					350					355		
gtg	cat	ctc	aag	gac	ctc	atc	tcc	ctg	tat	gaa	gcc	atg	cct	gac	tat	1219
Val	His	Leu	Lys	Asp	Leu	Ile	Ser	Leu	Tyr	Glu	Ala	Met	Pro	Asp	Tyr	
			360					365					370			
ctg	ggg	gac	ggg	aaa	gtg	aac	gtc	cat	aag	cta	ctg	gcc	cta	tac	aat	1267
Leu	Gly	Asp	Gly	Lys	Val	Asn	Val	His	Lys	Leu	Leu	Ala	Leu	Tyr	Asn	
		375					380					385				
cat	atc	agt	gaa	ttg	gtc	cag	ctg	caa	gag	gtg	gcc	cca	ccc	ttg	gag	1315
His	Ile	Ser	Glu	Leu	Val	Gln	Leu	Gln	Glu	Val	Ala	Pro	Pro	Leu	Glu	
	390					395					400					
gct	aac	aag	gac	ttg	gta	cac	ttg	ctg	acg	tta	tcc	ctg	gat	ctt	tac	1363
Ala	Asn	Lys	Asp	Leu	Val	His	Leu	Leu	Thr	Leu	Ser	Leu	Asp	Leu	Tyr	
405					410					415					420	
tac	act	gag	gat	gaa	atc	tat	gag	ctt	tcc	tat	gcc	cgg	gaa	cca	agg	1411
Tyr	Thr	Glu	Asp	Glu	Ile	Tyr	Glu	Leu	Ser	Tyr	Ala	Arg	Glu	Pro	Arg	
				425					430					435		
aac	cac	aga	gct	cca	cca	cta	aca	cct	tca	aag	cca	cca	gta	gta	gtg	1459
Asn	His	Arg	Ala	Pro	Pro	Leu	Thr	Pro	Ser	Lys	Pro	Pro	Val	Val	Val	
			440					445					450			
gac	tgg	gct	tct	gga	gtg	tct	ccc	aaa	cct	gat	cca	aaa	acc	att	agc	1507

Asp	Trp	Ala	Ser	Gly	Val	Ser	Pro	Lys	Pro	Asp	Pro	Lys	Thr	Ile	Ser	
		455					460					465				
aaa	cac	gtc	cag	agg	atg	gtg	gat	tct	gtc	ttc	aag	aac	tat	gat	cac	1555
Lys	His	Val	Gln	Arg	Met	Val	Asp	Ser	Val	Phe	Lys	Asn	Tyr	Asp	His	
		470				475					480					
gac	cag	gat	gga	tac	att	tct	cag	gaa	gaa	ttt	gaa	aag	att	gct	gcg	1603
Asp	Gln	Asp	Gly	Tyr	Ile	Ser	Gln	Glu	Glu	Phe	Glu	Lys	Ile	Ala	Ala	
					490					495					500	
agt	ttt	cca	ttt	tcc	ttc	tgt	gtg	atg	gac	aaa	gac	agg	gaa	ggc	ctc	1651
Ser	Phe	Pro	Phe	Ser	Phe	Cys	Val	Met	Asp	Lys	Asp	Arg	Glu	Gly	Leu	
				505					510					515		
atc	agc	agg	gat	gag	atc	aca	gcc	tac	ttc	atg	aga	gcc	agc	tca	atc	1699
Ile	Ser	Arg	Asp	Glu	Ile	Thr	Ala	Tyr	Phe	Met	Arg	Ala	Ser	Ser	Ile	
			520					525					530			
tat	tcc	aag	ctg	ggc	ctg	ggc	ttt	cct	cac	aac	ttc	caa	gag	acc	acc	1747
Tyr	Ser	Lys	Leu	Gly	Leu	Gly	Phe	Pro	His	Asn	Phe	Gln	Glu	Thr	Thr	
		535					540					545				
tac	ctg	aag	ccc	act	ttt	tgt	gac	aac	tgt	gct	gga	ttt	ctc	tggt	gga	1795
Tyr	Leu	Lys	Pro	Thr	Phe	Cys	Asp	Asn	Cys	Ala	Gly	Phe	Leu	Trp	Gly	
		550				555					560					
gtg	atc	aaa	caa	gga	tat	cga	tgt	aaa	gac	tgc	ggg	atg	aac	tgt	cac	1843
Val	Ile	Lys	Gln	Gly	Tyr	Arg	Cys	Lys	Asp	Cys	Gly	Met	Asn	Cys	His	
					570				575						580	
aaa	caa	tgc	aaa	gat	ctg	gtt	gtg	ttt	gag	tgt	aag	aag	cga	gcc	aag	1891
Lys	Gln	Cys	Lys	Asp	Leu	Val	Val	Phe	Glu	Cys	Lys	Lys	Arg	Ala	Lys	
				585					590					595		
aac	cca	gta	gct	ccc	aca	gag	aac	aac	act	tct	gtg	ggg	cca	gtg	tcc	1939
Asn	Pro	Val	Ala	Pro	Thr	Glu	Asn	Asn	Thr	Ser	Val	Gly	Pro	Val	Ser	
			600					605					610			
aac	ctt	tgc	tca	ttg	gga	gcc	aaa	gat	ctg	ctc	cat	gca	cct	gag	gaa	1987
Asn	Leu	Cys	Ser	Leu	Gly	Ala	Lys	Asp	Leu	Leu	His	Ala	Pro	Glu	Glu	
		615					620					625				
gga	cct	ttt	aca	ttc	cct	aat	ggg	gag	gct	gtg	gaa	cat	ggt	gag	gag	2035
Gly	Pro	Phe	Thr	Phe	Pro	Asn	Gly	Glu	Ala	Val	Glu	His	Gly	Glu	Glu	
		630				635					640					
agt	aag	gat	cgg	acc	atc	atg	ctg	atg	gga	gtg	tcc	tca	cag	aag	att	2083
Ser	Lys	Asp	Arg	Thr	Ile	Met	Leu	Met	Gly	Val	Ser	Ser	Gln	Lys	Ile	
					650					655					660	
tct	ctt	cgg	ctg	aag	agg	gct	gtt	gcc	cac	aag	gcc	acc	cag	act	gaa	2131
Ser	Leu	Arg	Leu	Lys	Arg	Ala	Val	Ala	His	Lys	Ala	Thr	Gln	Thr	Glu	
				665					670					675		
tca	cag	cct	tggt	att	ggc	agt	gag	ggc	cct	tca	ggt	ccc	ttt	gtg	ctg	2179
Ser	Gln	Pro	Trp	Ile	Gly	Ser	Glu	Gly	Pro	Ser	Gly	Pro	Phe	Val	Leu	
			680					685					690			
tct	tcc	cca	agg	aag	aca	gcc	cag	gat	act	cta	tat	gtg	ctt	ccc	agt	2227
Ser	Ser	Pro	Arg	Lys	Thr	Ala	Gln	Asp	Thr	Leu	Tyr	Val	Leu	Pro	Ser	
		695					700					705				
ccc	acc	tct	cca	tgt	cct	agc	cca	gtc	ttg	gtc	aga	aag	cgg	gct	ttt	2275
Pro	Thr	Ser	Pro	Cys	Pro	Ser	Pro	Val	Leu	Val	Arg	Lys	Arg	Ala	Phe	
		710				715					720					
gtc	aag	tggt	gag	aat	aaa	gac	tcc	ctc	ata	aaa	tca	aag	gag	gag	ctc	2323
Val	Lys	Trp	Glu	Asn	Lys	Asp	Ser	Leu	Ile	Lys	Ser	Lys	Glu	Glu	Leu	

725	730	735	740	
cgt cac ctc aga ctg cct acc tac caa gaa ctg gaa cag gaa ata aat	Arg His Leu Arg Leu Pro Thr Tyr Gln Glu Leu Glu Gln Glu Ile Asn	2371		
	745	750	755	
act ctg aaa gca gat aat gat gcc cta aag atc caa ctg aaa tat gca	Thr Leu Lys Ala Asp Asn Asp Ala Leu Lys Ile Gln Leu Lys Tyr Ala	2419		
	760	765	770	
cag aag aaa ata gaa tcc ctc cag ctt gaa aaa agc aat cat gtc tta	Gln Lys Lys Ile Glu Ser Leu Gln Leu Glu Lys Ser Asn His Val Leu	2467		
	775	780	785	
gct caa atg gag cag ggt gac tgt tct tagccagaa actaagtagc	Ala Gln Met Glu Gln Gly Asp Cys Ser	2514		
	790	795		
acaatctgta gatgagtata gtgatctcat ttccctaaact gtaatgcaca gacctgagga		2574		
actttacact gaccagcttt aaaacagtac tttaaaagga aaagcctgtt actgttttatt		2634		
tacctaaaag attcctaattg tgcagcactg ttttctcttt cagtttagttg actcaaaggg		2694		
ggaaaactaa agaatgcaaa acttttgcta ttcataccac tgatgttcat caaaagtatg		2754		
cgatctaaaa catgactatc atttcctgac aatggggcat ctcggtggcc tgcctggctg		2814		
attcttctctt aaaactaaaa tctctggaaa atggattttgc ttcttaccct gtgttttctg		2874		
caaactgact tactttgttc cagccaaagc ttgctaataa tagaaaacta cccatattcc		2934		
aaaagtagat ttccctctgta tcccagcata ctttgtgaac ctggctcctt cttcactacc		2994		
tcagatctaa tcaattagtc catgtaccct cttcctcac tacagtcata acacatgagc		3054		
atatctacct agaagccaat ttctactgat gtagccacaa ctttttagag gcctattaaa		3114		
catgacgtta tccaattgca ggtcaactta tagttgtagc cttacaactt acaggcatca		3174		
gaaaataagt aatcaaatta ggtacctgga aacatagcta ttaccatctc atattactgt		3234		
ctaattaaaa taaacataac gcaaacatgt ttgtccttat atattctaca gtggatagaa		3294		
ttaggaattg atggcttaaa aaaaaaagtc tatgaagagt ctgttttaact cttcatgttc		3354		
catctttctc ttctgaagta aactattttg aaagttctct ttttgaaatg aatttgtgct		3414		
taactgtctt cactattaat actatttaga aataagctaa ttggatcagt ggcttaaata		3474		
atagctgact gtgtgtacat atgtatataa tatgtatata caatatcagg catgcatgtg		3534		
gcttggaatt ttgtttcctc cataaaatgt ggaagtgaat taaacaagtt ttagtcattt		3594		
atacaaagtc acaaataata agttcagttt gtcacaagat taaattgctc acaaggtaaa		3654		
attgtattgt ttggcaaaat cacaagtaac aatcctgtga gttttctatt atgaaggtta		3714		
ataataaatg ggctcattta gttgcctggg cacctattca caaattcatt tgtcagcctc		3774		
tttttagttc tcttaaaaaa aaaaaaatca tatgatcatt ttcttttttg ggggtactta		3834		
gcttccatgc ctataaagtc tggtagcaga ctgacttgaa attcataaac aagttgtcca		3894		
attgccaaga atatgttaac aattaaaagt tccaaactaa agccaatagc accaagtctt		3954		
cataagaata caaagtatac atacagtatt gcttacctgg aggattcaga tcatttagga		4014		
attctctttg atgaaagatc agttcccatt tgagttcctc cttgcaactga gtttttagtga		4074		

GenBank accession number: U00180.1 (H. pylori strain NCTC 11637 genome)

tatagaacta gcttgtagtt agtgtttcat tacattataa agaatagttt tacacacgta 4134
 tttaccgttt tccaaattta aactcagaaa tacccaaagc aggccctgctt aagcccaacta 4194
 cctggcatat aaacttatag taacactttg ttacttttctt tttaatagga caagcatgag 4254
 ttaggacaaa ctctaaaaat tcatattctt cactattctt gttttccttt gattgatata 4314
 gaccaaagat ggtgtactct aattttttaa aacagtaatg gaacacaatt tttttcattc 4374
 ttcctcctct ccattcgaag taaagatccc cagtttagttt ttatataaat aatctatagg 4434
 gattcaaaag gtgtcacagt ccacttaatt agtcaaatta gcaatggcta aacagtatca 4494
 agtactgcag aatttatcac tgaaatggat aagaggaaat agtttagtca cagggttttta 4554
 cagtccagca agggccaaag aggtatagta tacaagttaa tagtatttgt gttgagcaac 4614
 atggggctag tgggatcaca gaaatctgga aaaaaaaaaa aaaaggcttt ggcttatcaa 4674
 gcctagtgtg aatttctgca tctcacacga ctttagtttg gccagggtatt tatctgccaa 4734
 aacaaggaca aatcttggtg tattaacagc aggggtcactt ctcattttct ttgctgactt 4794
 acctttttac tgaccgttgt gaatttctgt ctcaaatgt ataatataga aatgcaagaa 4854
 aaaaacaaat gtacagattg taaagttttt tgatacctaa tgtaagtttt ctttgtgtaa 4914
 tatttatatg ataaaagaca ttaggatccc tacaaaaaaaa aaaaaaaaaa aaaaaaaaaa 4974
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaggg atcgacaatt 5034
 acgactcccc gggggggccc ggtccaattc ccccaaaagg g 5075

<210> 8
 <211> 797
 <212> PRT
 <213> Homo sapiens

<400> 8
 Met Gly Thr Leu Gly Lys Ala Arg Glu Ala Pro Arg Lys Pro Ser His
 1 5 10 15
 Gly Cys Arg Ala Ala Ser Lys Ala Arg Leu Glu Ala Lys Pro Ala Asn
 20 25 30
 Ser Pro Phe Pro Ser His Pro Ser Leu Ala His Ile Thr Gln Phe Arg
 35 40 45
 Met Met Val Ser Leu Gly His Leu Ala Lys Gly Ala Ser Leu Asp Asp
 50 55 60
 Leu Ile Asp Ser Cys Ile Gln Ser Phe Asp Ala Asp Gly Asn Leu Cys
 65 70 75 80
 Arg Ser Asn Gln Leu Leu Gln Val Met Leu Thr Met His Arg Ile Val
 85 90 95
 Ile Ser Ser Ala Glu Leu Leu Gln Lys Val Ile Thr Leu Tyr Lys Asp
 100 105 110
 Ala Leu Ala Lys Asn Ser Pro Gly Leu Cys Leu Lys Ile Cys Tyr Phe
 115 120 125
 Val Arg Tyr Trp Ile Thr Glu Phe Trp Val Met Phe Lys Met Asp Ala
 130 135 140

Ser	Leu	Thr	Asp	Thr	Met	Glu	Glu	Phe	Gln	Glu	Leu	Val	Lys	Ala	Lys
145					150					155					160
Gly	Glu	Glu	Leu	His	Cys	Arg	Leu	Ile	Asp	Thr	Thr	Gln	Ile	Asn	Ala
				165					170					175	
Arg	Asp	Trp	Ser	Arg	Lys	Leu	Thr	Gln	Arg	Ile	Lys	Ser	Asn	Thr	Ser
			180					185					190		
Lys	Lys	Arg	Lys	Val	Ser	Leu	Leu	Phe	Asp	His	Leu	Glu	Pro	Glu	Glu
		195					200					205			
Leu	Ser	Glu	His	Leu	Thr	Tyr	Leu	Glu	Phe	Lys	Ser	Phe	Arg	Arg	Ile
	210					215					220				
Ser	Phe	Ser	Asp	Tyr	Gln	Asn	Tyr	Leu	Val	Asn	Ser	Cys	Val	Lys	Glu
225					230					235					240
Asn	Pro	Thr	Met	Glu	Arg	Ser	Ile	Ala	Leu	Cys	Asn	Gly	Ile	Ser	Gln
				245					250					255	
Trp	Val	Gln	Leu	Met	Val	Leu	Ser	Arg	Pro	Thr	Pro	Gln	Leu	Arg	Ala
			260					265					270		
Glu	Val	Phe	Ile	Lys	Phe	Ile	Gln	Val	Ala	Gln	Lys	Leu	His	Gln	Leu
		275					280					285			
Gln	Asn	Phe	Asn	Thr	Leu	Met	Ala	Val	Ile	Gly	Gly	Leu	Cys	His	Ser
	290					295					300				
Ser	Ile	Ser	Arg	Leu	Lys	Glu	Thr	Ser	Ser	His	Val	Pro	His	Glu	Ile
305					310					315					320
Asn	Lys	Val	Leu	Gly	Glu	Met	Thr	Glu	Leu	Leu	Ser	Ser	Ser	Arg	Asn
				325					330					335	
Tyr	Asp	Asn	Tyr	Arg	Arg	Ala	Tyr	Gly	Glu	Cys	Thr	Asp	Phe	Lys	Ile
			340					345					350		
Pro	Ile	Leu	Gly	Val	His	Leu	Lys	Asp	Leu	Ile	Ser	Leu	Tyr	Glu	Ala
		355					360					365			
Met	Pro	Asp	Tyr	Leu	Gly	Asp	Gly	Lys	Val	Asn	Val	His	Lys	Leu	Leu
	370					375					380				
Ala	Leu	Tyr	Asn	His	Ile	Ser	Glu	Leu	Val	Gln	Leu	Gln	Glu	Val	Ala
385					390					395					400
Pro	Pro	Leu	Glu	Ala	Asn	Lys	Asp	Leu	Val	His	Leu	Leu	Thr	Leu	Ser
				405					410					415	
Leu	Asp	Leu	Tyr	Tyr	Thr	Glu	Asp	Glu	Ile	Tyr	Glu	Leu	Ser	Tyr	Ala
			420					425					430		
Arg	Glu	Pro	Arg	Asn	His	Arg	Ala	Pro	Pro	Leu	Thr	Pro	Ser	Lys	Pro
		435					440					445			
Pro	Val	Val	Val	Asp	Trp	Ala	Ser	Gly	Val	Ser	Pro	Lys	Pro	Asp	Pro
	450					455					460				
Lys	Thr	Ile	Ser	Lys	His	Val	Gln	Arg	Met	Val	Asp	Ser	Val	Phe	Lys
465					470					475					480
Asn	Tyr	Asp	His	Asp	Gln	Asp	Gly	Tyr	Ile	Ser	Gln	Glu	Glu	Phe	Glu
				485					490					495	
Lys	Ile	Ala	Ala	Ser	Phe	Pro	Phe	Ser	Phe	Cys	Val	Met	Asp	Lys	Asp
			500					505					510		

Arg Glu Gly Leu Ile Ser Arg Asp Glu Ile Thr Ala Tyr Phe Met Arg
515 520 525

Ala Ser Ser Ile Tyr Ser Lys Leu Gly Leu Gly Phe Pro His Asn Phe
530 535 540

Gln Glu Thr Thr Tyr Leu Lys Pro Thr Phe Cys Asp Asn Cys Ala Gly
545 550 555 560

Phe Leu Trp Gly Val Ile Lys Gln Gly Tyr Arg Cys Lys Asp Cys Gly
565 570 575

Met Asn Cys His Lys Gln Cys Lys Asp Leu Val Val Phe Glu Cys Lys
580 585 590

Lys Arg Ala Lys Asn Pro Val Ala Pro Thr Glu Asn Asn Thr Ser Val
595 600 605

Gly Pro Val Ser Asn Leu Cys Ser Leu Gly Ala Lys Asp Leu Leu His
610 615 620

Ala Pro Glu Glu Gly Pro Phe Thr Phe Pro Asn Gly Glu Ala Val Glu
625 630 635 640

His Gly Glu Glu Ser Lys Asp Arg Thr Ile Met Leu Met Gly Val Ser
645 650 655

Ser Gln Lys Ile Ser Leu Arg Leu Lys Arg Ala Val Ala His Lys Ala
660 665 670

Thr Gln Thr Glu Ser Gln Pro Trp Ile Gly Ser Glu Gly Pro Ser Gly
675 680 685

Pro Phe Val Leu Ser Ser Pro Arg Lys Thr Ala Gln Asp Thr Leu Tyr
690 695 700

Val Leu Pro Ser Pro Thr Ser Pro Cys Pro Ser Pro Val Leu Val Arg
705 710 715 720

Lys Arg Ala Phe Val Lys Trp Glu Asn Lys Asp Ser Leu Ile Lys Ser
725 730 735

Lys Glu Glu Leu Arg His Leu Arg Leu Pro Thr Tyr Gln Glu Leu Glu
740 745 750

Gln Glu Ile Asn Thr Leu Lys Ala Asp Asn Asp Ala Leu Lys Ile Gln
755 760 765

Leu Lys Tyr Ala Gln Lys Lys Ile Glu Ser Leu Gln Leu Glu Lys Ser
770 775 780

Asn His Val Leu Ala Gln Met Glu Gln Gly Asp Cys Ser
785 790 795

<210> 9
<211> 3373
<212> DNA
<213> Rattus norvegicus

<220>
<221> CDS
<222> (197)..(2848)
<223> cAMP-GEFI

<400> 9
cgctaagcct ggggtggtgg taggaaggtc cagctcctcc ggccgatggc tgtccacttt 60
cctcccccca tcacaggtca gctggccagg tgagaaccac tggcaggtgg gccagctgt 120

ggtgagagat ccagctgtgg gggcaccgca ggtgcgaggt ctcccggacg tggttccgga															180	
gggcacgctg ctcaat atg gtg ctg aag aga atg cac cgt ccc cgg tgc tgc															232	
Met Val Leu Lys Arg Met His Arg Pro Arg Cys Cys																
1 5 10																
tct	tac	cag	cta	gtg	ttc	gag	cac	cgg	cgc	cca	agc	tgc	atc	cag	gga	280
Ser	Tyr	Gln	Leu	Val	Phe	Glu	His	Arg	Arg	Pro	Ser	Cys	Ile	Gln	Gly	
15 20 25																
ctt	cgc	tgg	acg	cca	ctt	acc	aac	agt	gag	ggc	tcc	ctg	gac	ttc	aga	328
Leu	Arg	Trp	Thr	Pro	Leu	Thr	Asn	Ser	Glu	Gly	Ser	Leu	Asp	Phe	Arg	
30 35 40																
gtg	agc	ctg	gag	cag	gcc	acc	aca	gag	cat	gtg	cac	aag	gcc	ggg	aag	376
Val	Ser	Leu	Glu	Gln	Ala	Thr	Thr	Glu	His	Val	His	Lys	Ala	Gly	Lys	
45 50 55 60																
ctc	ctg	tac	cgt	cat	ctc	ttg	gca	acg	tac	cct	acc	ctc	atc	cga	gac	424
Leu	Leu	Tyr	Arg	His	Leu	Leu	Ala	Thr	Tyr	Pro	Thr	Leu	Ile	Arg	Asp	
65 70 75																
aga	aaa	tac	cat	ctg	cga	cta	cat	cgg	cag	tgc	tgc	tct	ggc	cgg	gag	472
Arg	Lys	Tyr	His	Leu	Arg	Leu	His	Arg	Gln	Cys	Cys	Ser	Gly	Arg	Glu	
80 85 90																
cta	gtg	gat	ggg	atc	ttg	gct	ctg	ggt	ctt	ggg	gtc	cac	tca	cgg	agc	520
Leu	Val	Asp	Gly	Ile	Leu	Ala	Leu	Gly	Leu	Gly	Val	His	Ser	Arg	Ser	
95 100 105																
caa	gct	gtg	ggc	atc	tgc	cag	gtg	ttg	ctg	gat	gag	ggt	gcc	ctt	tgc	568
Gln	Ala	Val	Gly	Ile	Cys	Gln	Val	Leu	Leu	Asp	Glu	Gly	Ala	Leu	Cys	
110 115 120																
cat	gta	aaa	cat	gac	tgg	acc	ttc	cag	gac	cga	gac	gcc	caa	ttc	tac	616
His	Val	Lys	His	Asp	Trp	Thr	Phe	Gln	Asp	Arg	Asp	Ala	Gln	Phe	Tyr	
125 130 135 140																
aga	ttc	cct	gga	cgg	gag	ccc	cag	cct	gca	gga	act	cat	gac	gtg	gaa	664
Arg	Phe	Pro	Gly	Pro	Glu	Pro	Gln	Pro	Ala	Gly	Thr	His	Asp	Val	Glu	
145 150 155																
gag	gag	ctt	gtt	gag	gca	atg	gcc	cta	ctc	tcc	cag	cga	ggg	cct	gat	712
Glu	Glu	Leu	Val	Glu	Ala	Met	Ala	Leu	Leu	Ser	Gln	Arg	Gly	Pro	Asp	
160 165 170																
gcc	cta	ctc	act	gtt	gca	ctc	cgg	aag	tcc	cgg	ggt	cag	cgt	aca	gat	760
Ala	Leu	Leu	Thr	Val	Ala	Leu	Arg	Lys	Ser	Pro	Gly	Gln	Arg	Thr	Asp	
175 180 185																
gaa	gag	ctg	gac	ctc	atc	ttc	gag	gag	ctc	gta	cat	atc	aag	gcg	gtg	808
Glu	Glu	Leu	Asp	Leu	Ile	Phe	Glu	Glu	Leu	Val	His	Ile	Lys	Ala	Val	
190 195 200																
gct	cac	ctt	tct	aac	tcg	gtg	aaa	cgg	gaa	cta	gct	gct	gtt	ctg	ctc	856
Ala	His	Leu	Ser	Asn	Ser	Val	Lys	Arg	Glu	Leu	Ala	Ala	Val	Leu	Leu	
205 210 215 220																
ttt	gaa	cca	cac	agc	aag	gca	gga	act	gtg	ttg	ttc	agc	cag	ggg	gac	904
Phe	Glu	Pro	His	Ser	Lys	Ala	Gly	Thr	Val	Leu	Phe	Ser	Gln	Gly	Asp	
225 230 235																
aag	ggt	acc	tca	tgg	tac	att	atc	tgg	aag	gga	tct	gtc	aat	gtg	gtg	952
Lys	Gly	Thr	Ser	Trp	Tyr	Ile</										

	255					260					265					
gga Gly	cag Gln 270	ctg Leu	gct Ala	ctg Leu	gtg Val	aac Asn 275	gac Asp	gca Ala	cct Pro	cga Arg	gca Ala 280	gcc Ala	acc Thr	atc Ile	atc Ile	1048
ctt Leu 285	cga Arg	gaa Glu	aat Asn	aac Asn	tgt Cys 290	cac His	ttt Phe	ctg Leu	cgt Arg	gtg Val 295	gac Asp	aag Lys	cag Gln	gac Asp	ttc Phe 300	1096
aac Asn	cgc Arg	atc Ile	atc Ile	aag Lys 305	gat Asp	gtg Val	gaa Glu	gca Ala	aaa Lys 310	acc Thr	atg Met	aga Arg	ctg Leu	gaa Glu 315	gaa Glu	1144
cac His	ggc Gly	aaa Lys	gtg Val 320	gtg Val	tta Leu	gtt Val	ttg Leu	gag Glu 325	aga Arg	acc Thr	tct Ser	cag Gln	ggt Gly 330	gct Ala	ggc Gly	1192
cct Pro	tcc Ser	cgc Arg 335	cct Pro	ccg Pro	acc Thr	cca Pro	ggc Gly 340	agg Arg	aac Asn	cga Arg	tat Tyr	acg Thr 345	gta Val	atg Met	tct Ser	1240
ggc Gly	acc Thr 350	cca Pro	gag Glu	aaa Lys	atc Ile	cta Leu 355	gaa Glu	ctc Leu	ctg Leu	ttg Leu	gag Glu 360	gct Ala	atg Met	aga Arg	ccg Pro	1288
gat Asp 365	tcc Ser	agt Ser	gct Ala	cat His	gac Asp 370	cca Pro	aca Thr	gag Glu	aca Thr	ttc Phe 375	ctc Leu	agt Ser	gac Asp	ttc Phe	ctg Leu 380	1336
ctg Leu	acg Thr	cac His	agt Ser	gtc Val 385	ttc Phe	atg Met	ccc Pro	tgc Cys	aca Thr 390	cag Gln	ctc Leu	ttt Phe	gcc Ala	gcc Ala 395	ctc Leu	1384
ctg Leu	cac His	cac His	ttc Phe 400	cac His	gtg Val	gag Glu	cca Pro	tca Ser 405	gag Glu	cct Pro	gcc Ala	ggg Gly 410	ggc Gly 410	agc Ser	gag Glu	1432
cag Gln	gaa Glu	cgc Arg 415	agc Ser	acc Thr	tac Tyr	atc Ile	tgc Cys 420	aac Asn	aag Lys	agg Arg	cag Gln	cag Gln 425	att Ile	ctg Leu	cgt Arg	1480
ctg Leu	gtc Val 430	agc Ser	cgg Arg	tgg Trp	gtg Val	gcc Ala 435	ctc Leu	tac Tyr	agc Ser	ccc Pro	atg Met 440	ctc Leu	cgc Arg	tca Ser	gat Asp	1528
ccc Pro 445	gtg Val	gcc Ala	acc Thr	agc Ser	ttc Phe 450	ctc Leu	cag Gln	aaa Lys	ctc Leu	tca Ser 455	gac Asp	ctg Leu	gtg Val	agc Ser	aga Arg 460	1576
gat Asp	acc Thr	cga Arg	ctt Leu	agc Ser 465	aac Asn	ttg Leu	ctg Leu	agg Arg	gaa Glu 470	cag Gln	tat Tyr	ccg Pro	gag Glu	aga Arg 475	ccg Arg	1624
cga Arg	cac His	cac His	agg Arg 480	ttg Leu	gag Glu	aat Asn	ggc Gly	tgt Cys 485	ggg Gly	aat Asn	gta Val	tct Ser	cct Pro 490	cag Gln	acc Thr	1672
aag Lys	gcc Ala	cgg Arg 495	aat Asn	gca Ala	cct Pro	gtt Val	tgg Trp 500	ttt Phe	cct Pro	aac Asn	cat His	gag Glu 505	gaa Glu	ccc Pro	ctc Leu	1720
coa Pro	agc Ser 510	agc Ser	gct Ala	ggg Gly	gcc Ala	atc Ile 515	cga Arg	gtc Val	ggg Gly	gac Asp	aaa Lys 520	gtc Val	ccc Pro	tat Tyr	gat Asp	1768
atc Ile 525	tgc Cys	aga Arg	ccc Pro	gac Asp	cac His 530	tcg Ser	gtg Val	ctg Leu	acc Thr	ctg Leu 535	cac His	ctg Leu	ccg Pro	gtg Val	aca Thr 540	1816

gcc Ala	tcc Ser	gtg Val	agg Arg	gaa Glu 545	gtg Val	atg Met	gca Ala	gct Ala	ttg Leu 550	gcc Ala	cat His	gag Glu	gac Asp	cac His 555	tgg Trp	1864
acc Thr	aag Lys	ggg Gly	cag Gln 560	gtg Val	ctg Leu	gta Val	aag Lys	gtc Val 565	aat Asn	tct Ser	gcc Ala	ggt Gly	gat Asp 570	gtc Val	gtt Val	1912
ggc Gly	ttg Leu	cag Gln 575	cca Pro	gat Asp	gcc Ala	cgc Arg	ggt Gly 580	gtg Val	gcc Ala	aca Thr	tcc Ser	ctg Leu 585	ggg Gly	ctc Leu	aat Asn	1960
gag Glu	cgg Arg 590	atc Ile	ttt Phe	gtt Val	gtc Val	gac Asp 595	cca Pro	cag Gln	gaa Glu	gtg Val	cac His 600	gag Glu	ctg Leu	acc Thr	cca Pro	2008
cac His 605	cct Pro	gag Glu	cag Gln	ctg Leu	ggg Gly 610	ccc Pro	act Thr	ctg Leu	ggt Gly	tct Ser 615	tct Ser	gag Glu	atg Met	ctg Leu	gac Asp 620	2056
cta Leu	gtg Val	agt Ser	gcc Ala	aag Lys 625	gac Asp	ctg Leu	gca Ala	ggc Gly	cag Gln 630	ctc Leu	aca Thr	gag Glu	cat His	gac Asp 635	tgg Trp	2104
aac Asn	ctc Leu	ttc Phe	aac Asn 640	agg Arg	atc Ile	cac His	cag Gln	gtg Val 645	gag Glu	ctg Leu	atc Ile	cac His	tat Tyr 650	gta Val	ctg Leu	2152
ggc Gly	ccc Pro	cag Gln 655	cac His	ctg Leu	cgg Arg	gac Asp	gtc Val 660	acc Thr	act Thr	gca Ala	aac Asn	ctg Leu 665	gag Glu	cgc Arg	ttc Phe	2200
atg Met	cga Arg 670	cgc Arg	ttc Phe	aac Asn	gag Glu	ctg Leu 675	cag Gln	tac Tyr	tgg Trp	gtg Val	gcc Ala 680	acc Thr	gaa Glu	ctc Leu	tgt Cys	2248
ctc Leu 685	tgc Cys	ccg Pro	gtt Val	cct Pro	ggc Gly 690	ccc Pro	cgg Arg	gct Ala	cag Gln 695	cta Leu	ctc Leu	cgg Arg	aag Lys	ttc Phe	atc Ile 700	2296
aag Lys	ctg Leu	gca Ala	gcc Ala	cac His 705	ctc Leu	aag Lys	gag Glu	cag Gln	aag Lys 710	aat Asn	ctc Leu	aac Asn	tct Ser	ttc Phe 715	ttt Phe	2344
gcg Ala	gtc Val	atg Met	ttt Phe 720	ggc Gly	ctc Leu	agc Ser	aac Asn	tcg Ser 725	gcc Ala	atc Ile	agc Ser	cgc Arg	ctg Leu 730	gcc Ala	cac His	2392
acc Thr	tgg Trp	gag Glu 735	cgt Arg	ctg Leu	ccc Pro	cat His	aaa Lys 740	gta Val	cgg Arg	aag Lys	ctg Leu	tac Tyr 745	tcg Ser	gcc Ala	ctg Leu	2440
gaa Glu 750	agg Arg	ttg Leu	ctg Leu	gac Asp	cct Pro	tcc Ser 755	tgg Trp	aac Asn	cac His	cga Arg	gtg Val 760	tac Tyr	cga Arg	ttg Leu	gct Ala	2488
ctc Leu 765	acc Thr	aag Lys	ctc Leu	tct Ser	cct Pro 770	cct Pro	gtc Val	atc Ile	cct Pro 775	ttc Phe	atg Met	ccc Pro	ctg Leu	cta Leu	ctc Leu 780	2536
aaa Lys	gac Asp	atg Met	acc Thr	ttc Phe 785	att Ile	cat His	gaa Glu	ggg Gly	aac Asn 790	cac His	aca Thr	ctg Leu	gta Val	gaa Glu 795	aac Asn	2584
ctc Leu	atc Ile	aac Asn	ttt Phe 800	gag Glu	aag Lys	atg Met	cga Arg	atg Met 805	atg Met	gcc Ala	aga Arg	gcc Ala	gtg Val 810	cgg Arg	atg Met	2632

ctc cac cac tgc cga agc cac agc acc gcg cct cta tca cca ctc aga 2680
 Leu His His Cys Arg Ser His Ser Thr Ala Pro Leu Ser Pro Leu Arg
 815 820 825

agc cgg gtt tcc cac atc cac gag gac agc cag gca tca aga atc tcc 2728
 Ser Arg Val Ser His Ile His Glu Asp Ser Gln Ala Ser Arg Ile Ser
 830 835 840

aca tgt tcc gag cag tcc ctg agc acc cgg agt cca gcc agc acc tgg 2776
 Thr Cys Ser Glu Gln Ser Leu Ser Thr Arg Ser Pro Ala Ser Thr Trp
 845 850 855 860

gct tat gtc cag cag ctg aag gtc att gac aac cag cgg gaa ctg tcc 2824
 Ala Tyr Val Gln Gln Leu Lys Val Ile Asp Asn Gln Arg Glu Leu Ser
 865 870 875

cgc ctc tcc cgg gaa ctg gaa cca tgaggaagga ctggctggag caggcattc 2878
 Arg Leu Ser Arg Glu Leu Glu Pro
 880

tctcagagaa agccagagcc tgtgcaacca agaggtccag aggccagcca cagctgggca 2938

gggctctcca cagagcggac tcaaggccct ggagtgggca gtgtcgaggc agctgtcctc 2998

tgtgatgact gtcagctgtg aagatctttg atgttcacgg ccaaggaaaa ggggccattg 3058

aggccccaga ggggtaggag agctgggagg tgcaggactc tggttcagta gagagccctc 3118

ccaccgggct ttctgcatgt ctgtatgtct gtacatgcag ctgtgtgtcc tggatgccag 3178

gccctgtgct tgtattcaca ggagccagcg agctcacatc tgcattggtgt gtgtgtgtgt 3238

gtgtgtgtgt gtgtgtgtgt gtgtgtgtga gcaactgggtg gtgccttgcc tgggagaggg 3298

tggggagttc tgctattctc accacacatc tgagataaac agctgggtgt gggcaaaaaa 3358

aaaaaaaaa aaaaaa 3373

<210> 10
 <211> 884
 <212> PRT
 <213> Rattus norvegicus

<400> 10
 Met Val Leu Lys Arg Met His Arg Pro Arg Cys Cys Ser Tyr Gln Leu
 1 5 10 15

Val Phe Glu His Arg Arg Pro Ser Cys Ile Gln Gly Leu Arg Trp Thr
 20 25 30

Pro Leu Thr Asn Ser Glu Gly Ser Leu Asp Phe Arg Val Ser Leu Glu
 35 40 45

Gln Ala Thr Thr Glu His Val His Lys Ala Gly Lys Leu Leu Tyr Arg
 50 55 60

His Leu Leu Ala Thr Tyr Pro Thr Leu Ile Arg Asp Arg Lys Tyr His
 65 70 75 80

Leu Arg Leu His Arg Gln Cys Cys Ser Gly Arg Glu Leu Val Asp Gly
 85 90 95

Ile Leu Ala Leu Gly Leu Gly Val His Ser Arg Ser Gln Ala Val Gly
 100 105 110

Ile Cys Gln Val Leu Leu Asp Glu Gly Ala Leu Cys His Val Lys His
 115 120 125

Asp	Trp	Thr	Phe	Gln	Asp	Arg	Asp	Ala	Gln	Phe	Tyr	Arg	Phe	Pro	Gly
130						135					140				
Pro	Glu	Pro	Gln	Pro	Ala	Gly	Thr	His	Asp	Val	Glu	Glu	Glu	Leu	Val
145					150					155					160
Glu	Ala	Met	Ala	Leu	Leu	Ser	Gln	Arg	Gly	Pro	Asp	Ala	Leu	Leu	Thr
				165					170					175	
Val	Ala	Leu	Arg	Lys	Ser	Pro	Gly	Gln	Arg	Thr	Asp	Glu	Glu	Leu	Asp
			180					185					190		
Leu	Ile	Phe	Glu	Glu	Leu	Val	His	Ile	Lys	Ala	Val	Ala	His	Leu	Ser
		195					200					205			
Asn	Ser	Val	Lys	Arg	Glu	Leu	Ala	Ala	Val	Leu	Leu	Phe	Glu	Pro	His
	210					215					220				
Ser	Lys	Ala	Gly	Thr	Val	Leu	Phe	Ser	Gln	Gly	Asp	Lys	Gly	Thr	Ser
225					230					235					240
Trp	Tyr	Ile	Ile	Trp	Lys	Gly	Ser	Val	Asn	Val	Val	Thr	Arg	Gly	Lys
				245					250					255	
Gly	Leu	Val	Thr	Thr	Leu	His	Glu	Gly	Asp	Asp	Phe	Gly	Gln	Leu	Ala
			260					265					270		
Leu	Val	Asn	Asp	Ala	Pro	Arg	Ala	Ala	Thr	Ile	Ile	Leu	Arg	Glu	Asn
		275					280					285			
Asn	Cys	His	Phe	Leu	Arg	Val	Asp	Lys	Gln	Asp	Phe	Asn	Arg	Ile	Ile
	290					295					300				
Lys	Asp	Val	Glu	Ala	Lys	Thr	Met	Arg	Leu	Glu	Glu	His	Gly	Lys	Val
305					310					315					320
Val	Leu	Val	Leu	Glu	Arg	Thr	Ser	Gln	Gly	Ala	Gly	Pro	Ser	Arg	Pro
				325					330					335	
Pro	Thr	Pro	Gly	Arg	Asn	Arg	Tyr	Thr	Val	Met	Ser	Gly	Thr	Pro	Glu
			340					345					350		
Lys	Ile	Leu	Glu	Leu	Leu	Leu	Glu	Ala	Met	Arg	Pro	Asp	Ser	Ser	Ala
		355					360					365			
His	Asp	Pro	Thr	Glu	Thr	Phe	Leu	Ser	Asp	Phe	Leu	Leu	Thr	His	Ser
	370					375					380				
Val	Phe	Met	Pro	Cys	Thr	Gln	Leu	Phe	Ala	Ala	Leu	Leu	His	His	Phe
385					390					395					400
His	Val	Glu	Pro	Ser	Glu	Pro	Ala	Gly	Gly	Ser	Glu	Gln	Glu	Arg	Ser
				405					410					415	
Thr	Tyr	Ile	Cys	Asn	Lys	Arg	Gln	Gln	Ile	Leu	Arg	Leu	Val	Ser	Arg
			420					425					430		
Trp	Val	Ala	Leu	Tyr	Ser	Pro	Met	Leu	Arg	Ser	Asp	Pro	Val	Ala	Thr
		435					440					445			
Ser	Phe	Leu	Gln	Lys	Leu	Ser	Asp	Leu	Val	Ser	Arg	Asp	Thr	Arg	Leu
	450					455					460				
Ser	Asn	Leu	Leu	Arg	Glu	Gln	Tyr	Pro	Glu	Arg	Arg	Arg	His	His	Arg
465					470					475					480
Leu	Glu	Asn	Gly	Cys	Gly	Asn	Val	Ser	Pro	Gln	Thr	Lys	Ala	Arg	Asn
				485					490					495	

Ala	Pro	Val	Trp	Phe	Pro	Asn	His	Glu	Glu	Pro	Leu	Pro	Ser	Ser	Ala
			500					505					510		
Gly	Ala	Ile	Arg	Val	Gly	Asp	Lys	Val	Pro	Tyr	Asp	Ile	Cys	Arg	Pro
		515					520					525			
Asp	His	Ser	Val	Leu	Thr	Leu	His	Leu	Pro	Val	Thr	Ala	Ser	Val	Arg
	530					535					540				
Glu	Val	Met	Ala	Ala	Leu	Ala	His	Glu	Asp	His	Trp	Thr	Lys	Gly	Gln
545					550					555					560
Val	Leu	Val	Lys	Val	Asn	Ser	Ala	Gly	Asp	Val	Val	Gly	Leu	Gln	Pro
				565					570					575	
Asp	Ala	Arg	Gly	Val	Ala	Thr	Ser	Leu	Gly	Leu	Asn	Glu	Arg	Ile	Phe
			580					585					590		
Val	Val	Asp	Pro	Gln	Glu	Val	His	Glu	Leu	Thr	Pro	His	Pro	Glu	Gln
		595					600					605			
Leu	Gly	Pro	Thr	Leu	Gly	Ser	Ser	Glu	Met	Leu	Asp	Leu	Val	Ser	Ala
	610					615					620				
Lys	Asp	Leu	Ala	Gly	Gln	Leu	Thr	Glu	His	Asp	Trp	Asn	Leu	Phe	Asn
625					630					635					640
Arg	Ile	His	Gln	Val	Glu	Leu	Ile	His	Tyr	Val	Leu	Gly	Pro	Gln	His
				645					650					655	
Leu	Arg	Asp	Val	Thr	Thr	Ala	Asn	Leu	Glu	Arg	Phe	Met	Arg	Arg	Phe
			660					665					670		
Asn	Glu	Leu	Gln	Tyr	Trp	Val	Ala	Thr	Glu	Leu	Cys	Leu	Cys	Pro	Val
		675					680					685			
Pro	Gly	Pro	Arg	Ala	Gln	Leu	Leu	Arg	Lys	Phe	Ile	Lys	Leu	Ala	Ala
	690					695					700				
His	Leu	Lys	Glu	Gln	Lys	Asn	Leu	Asn	Ser	Phe	Phe	Ala	Val	Met	Phe
705					710					715					720
Gly	Leu	Ser	Asn	Ser	Ala	Ile	Ser	Arg	Leu	Ala	His	Thr	Trp	Glu	Arg
				725					730					735	
Leu	Pro	His	Lys	Val	Arg	Lys	Leu	Tyr	Ser	Ala	Leu	Glu	Arg	Leu	Leu
			740					745					750		
Asp	Pro	Ser	Trp	Asn	His	Arg	Val	Tyr	Arg	Leu	Ala	Leu	Thr	Lys	Leu
		755					760					765			
Ser	Pro	Pro	Val	Ile	Pro	Phe	Met	Pro	Leu	Leu	Leu	Lys	Asp	Met	Thr
	770					775					780				
Phe	Ile	His	Glu	Gly	Asn	His	Thr	Leu	Val	Glu	Asn	Leu	Ile	Asn	Phe
785					790					795					800
Glu	Lys	Met	Arg	Met	Met	Ala	Arg	Ala	Val	Arg	Met	Leu	His	His	Cys
				805					810					815	
Arg	Ser	His	Ser	Thr	Ala	Pro	Leu	Ser	Pro	Leu	Arg	Ser	Arg	Val	Ser
			820					825					830		
His	Ile	His	Glu	Asp	Ser	Gln	Ala	Ser	Arg	Ile	Ser	Thr	Cys	Ser	Glu
		835					840					845			
Gln	Ser	Leu	Ser	Thr	Arg	Ser	Pro	Ala	Ser	Thr	Trp	Ala	Tyr	Val	Gln
	850					855					860				

Gln Leu Lys Val Ile Asp Asn Gln Arg Glu Leu Ser Arg Leu Ser Arg
865 870 875 880

Glu Leu Glu Pro

<210> 11
<211> 3394
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (216)..(2858)
<223> cAMP-GEFI

<400> 11
ggatccccctt atcaaagctg atggggggtcg ctgggggaccc cccagccttt tgctagagcc 60
tgcacggctg tgggagcttg aaaagaaaca tgaaggtggg ctggccaggt gagagctgct 120
ggcaggtggg cctggctgtg gaggatagcc cagctctggg agcaccgagg gtgggagccc 180
tccctgacgt ggtgccggag gggacactac tcaac atg gtg ttg aga agg atg 233
Met Val Leu Arg Arg Met
1 5
cac cgg ccc cga agc tgc tcc tac cag ctg ctg ctg gag cac cag cat 281
His Arg Pro Arg 10 Ser Cys Ser Tyr Gln Leu Leu Leu Glu His Gln His 20
ccg agc tgc atc cag ggg ctg cgc tgg aca cca ctc acc aac agc gag 329
Pro Ser Cys 25 Ile Gln Gly Leu Arg 30 Trp Thr Pro Leu Thr 35 Asn Ser Glu
gag tcc ctg gat ttc agc gag agc ctg gag cag gcc tcc aca gag cgg 377
Glu Ser 40 Leu Asp Phe Ser Glu 45 Ser Leu Glu Gln 50 Ala Ser Thr Glu Arg
gtg ctc agg gct ggg agg cag ctg cat cag cat cta ctg gcc acc tgc 425
Val 55 Leu Arg Ala Gly Arg 60 Gln Leu His Gln His 65 Leu Leu Ala Thr Cys 70
cca aac ctc atc cga gac cgg aag tac cac ctt agg ctc tat cgg cag 473
Pro Asn Leu Ile Arg 75 Asp Arg Lys Tyr His 80 Leu Arg Leu Tyr Arg 85 Gln
tgc tgc tct ggc cgg gag ctg gtg gat ggg atc ttg gcc ctg gga ctt 521
Cys Cys Ser Gly 90 Arg Glu Leu Val Asp 95 Gly Ile Leu Ala Leu Gly Leu 100
ggg gtc cat tcc cgg agc caa gtt gtg gga atc tgc cag gtg ctg ctg 569
Gly Val 105 Ser Arg Ser Gln Val Val Gly Ile Cys 115 Gln Val Leu Leu
gat gaa ggt gcc ctc tgc cat gtg aaa cac gac tgg gcc ttc cag gac 617
Asp Glu Gly Ala Leu Cys 125 Val Lys His Asp Trp 130 Ala Phe Gln Asp
cga gat gcc caa ttc tac cgg ttc ccc ggg ccc gag ccc gag ccc gtg 665
Arg Asp Ala Gln Phe 140 Arg Phe Pro Gly Pro 145 Glu Pro Glu Pro Val 150
gga act cat gag atg gag gag gag ttg gcc gaa gct gtg gcc ctg ctc 713
Gly Thr His Glu Met 155 Glu Glu Glu Leu 160 Ala Glu Ala Val Ala Leu Leu 165
tcc cag cgg ggg cct gac gcc ctg ctc act gtg gca ctt cga aag ccc 761

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Ser	Gln	Arg	Gly	Pro	Asp	Ala	Leu	Leu	Thr	Val	Ala	Leu	Arg	Lys	Pro	
			170					175					180			
cca	ggt	cag	cgc	acg	gat	gaa	gag	ctg	gac	ctc	atc	ttt	gag	gag	ctg	809
Pro	Gly	Gln	Arg	Thr	Asp	Glu	Glu	Leu	Asp	Leu	Ile	Phe	Glu	Glu	Leu	
		185					190					195				
ctg	cac	atc	aag	gct	gtg	gcc	cac	ctc	tcc	aac	tcg	gtg	aag	cga	gaa	857
Leu	His	Ile	Lys	Ala	Val	Ala	His	Leu	Ser	Asn	Ser	Val	Lys	Arg	Glu	
	200					205					210					
tta	gcg	gct	gtt	ctg	ctc	ttt	gaa	cca	cac	agc	aag	gca	ggg	acc	gtg	905
Leu	Ala	Ala	Val	Leu	Leu	Phe	Glu	Pro	His	Ser	Lys	Ala	Gly	Thr	Val	
215				220						225					230	
ttg	ttc	agc	cag	ggg	gac	aag	ggc	act	tcg	tgg	tac	att	atc	tgg	aag	953
Leu	Phe	Ser	Gln	Gly	Asp	Lys	Gly	Thr	Ser	Trp	Tyr	Ile	Ile	Trp	Lys	
			235						240					245		
gga	tct	gtc	aac	gtg	gtg	acc	cat	ggc	aag	ggg	ctg	gtg	acc	acc	ctg	1001
Gly	Ser	Val	Asn	Val	Val	Thr	His	Gly	Lys	Gly	Leu	Val	Thr	Thr	Leu	
			250					255					260			
cat	gag	gga	gat	gat	ttt	gga	cag	ctg	gct	ctg	gtg	aat	gat	gca	ccc	1049
His	Glu	Gly	Asp	Asp	Phe	Gly	Gln	Leu	Ala	Leu	Val	Asn	Asp	Ala	Pro	
		265					270					275				
cgg	gca	gcc	acc	atc	atc	ctg	cga	gaa	tac	aac	tgt	cat	ttc	ctg	cgt	1097
Arg	Ala	Ala	Thr	Ile	Ile	Leu	Arg	Glu	Tyr	Asn	Cys	His	Phe	Leu	Arg	
	280					285					290					
gtg	gac	aag	cag	gac	ttc	aac	cgt	atc	atc	aag	gat	gtg	gag	gca	aag	1145
Val	Asp	Lys	Gln	Asp	Phe	Asn	Arg	Ile	Ile	Lys	Asp	Val	Glu	Ala	Lys	
295				300						305					310	
acc	atg	cgg	ctg	gaa	gaa	cat	ggc	aaa	gtg	gtg	ctg	gtg	ctg	gag	aga	1193
Thr	Met	Arg	Leu	Glu	Glu	His	Gly	Lys	Val	Val	Leu	Val	Leu	Glu	Arg	
			315						320					325		
gcc	tct	cag	ggc	gcc	ggc	cct	tcc	cga	ccc	cca	acc	cca	ggc	agg	aac	1241
Ala	Ser	Gln	Gly	Ala	Gly	Pro	Ser	Arg	Pro	Pro	Thr	Pro	Gly	Arg	Asn	
			330					335					340			
cgg	tat	aca	gtg	atg	tct	ggc	act	cca	gat	aag	atc	cta	gag	ctt	ctg	1289
Arg	Tyr	Thr	Val	Met	Ser	Gly	Thr	Pro	Asp	Lys	Ile	Leu	Glu	Leu	Leu	
		345					350					355				
ttg	gag	gcc	atg	gga	cta	gat	tcc	agt	gct	cat	gac	cca	aaa	gaa	aca	1337
Leu	Glu	Ala	Met	Gly	Leu	Asp	Ser	Ser	Ala	His	Asp	Pro	Lys	Glu	Thr	
	360					365					370					
ttc	ctc	agc	gac	ttc	ctc	ctg	acc	cac	agg	gtc	ttc	atg	ccc	agc	gcc	1385
Phe	Leu	Ser	Asp	Phe	Leu	Leu	Thr	His	Arg	Val	Phe	Met	Pro	Ser	Ala	
375					380					385					390	
caa	ctc	tgc	gct	gcc	ctt	ctg	cac	cac	ttc	cat	gtg	gag	cct	gcg	ggt	1433
Gln	Leu	Cys	Ala	Ala	Leu	Leu	His	His	Phe	His	Val	Glu	Pro	Ala	Gly	
				395					400					405		
ggc	agc	gag	cag	gag	cgc	agc	acc	tac	gtc	tgc	aac	aag	agg	cag	cag	1481
Gly	Ser	Glu	Gln	Glu	Arg	Ser	Thr	Tyr	Val	Cys	Asn	Lys	Arg	Gln	Gln	
			410					415					420			
atc	ttg	cgg	ctg	gtc	agc	cag	tgg	gtg	gcc	ctg	tat	ggc	tcc	atg	ctc	1529
Ile	Leu	Arg	Leu	Val	Ser	Gln	Trp	Val	Ala	Leu	Tyr	Gly	Ser	Met	Leu	
		425					430					435				
cac	act	gac	cct	gtg	gcc	acc	agc	ttc	ctc	cac	aaa	ctc	tca	gac	ctg	1577
His	Thr	Asp	Pro	Val	Ala	Thr	Ser	Phe	Leu	His	Lys	Leu	Ser	Asp	Leu	

440	445	450	
gtg ggc agg gac acc cga ctc agc aac ctg ctg agg gag cag tgg cca Val Gly Arg Asp Thr Arg Leu Ser Asn Leu Leu Arg Glu Gln Trp Pro 455 460 465 470			1625
gag agg cgg cga tgc cac agg ttg gag aat ggc tgt ggg aat gca tct Glu Arg Arg Arg Cys His Arg Leu Glu Asn Gly Cys Gly Asn Ala Ser 475 480 485			1673
cct cag atg aag gcc cgg aac ttg cct gtt tgg ctc ccc aac cag gac Pro Gln Met Lys Ala Arg Asn Leu Pro Val Trp Leu Pro Asn Gln Asp 490 500			1721
gag ccc ctt cct ggc agc agc tgt gcc atc caa gtt ggg gat aaa gtc Glu Pro Leu Pro Gly Ser Ser Cys Ala Ile Gln Val Gly Asp Lys Val 505 510 515			1769
ccc tat gac atc tgc cgg cca gac cac tca gtg ttg acc ctg cag ctg Pro Tyr Asp Ile Cys Arg Pro Asp His Ser Val Leu Thr Leu Gln Leu 520 525 530			1817
cct gtg aca gcc tcc gtg aga gag gtg atg gca gcg ttg gcc cag gag Pro Val Thr Ala Ser Val Arg Glu Val Met Ala Ala Leu Ala Gln Glu 535 540 545 550			1865
gat ggc tgg acc aag ggg cag gtg ctg gtg aag gtc aat tct gca ggt Asp Gly Trp Thr Lys Gly Gln Val Leu Val Lys Val Asn Ser Ala Gly 555 560 565			1913
gat gcc att ggc ctg cag cca gat gcc cgt ggt gtg gcc aca tct ctg Asp Ala Ile Gly Leu Gln Pro Asp Ala Arg Gly Val Ala Thr Ser Leu 570 575 580			1961
ggg ctc aat gag cgt ctc ttt gtt gtc aac cca cag gaa gtg cat gag Gly Leu Asn Glu Arg Leu Phe Val Val Asn Pro Gln Glu Val His Glu 585 590 595			2009
ctg atc cca cac cct gac cag ctg ggg ccc act gtg ggc tct gct gag Leu Ile Pro His Pro Asp Gln Leu Gly Pro Thr Val Gly Ser Ala Glu 600 605 610			2057
ggg ctg gac ctg gtg agt gcc aag gac ctg gca ggc cag ctg acg gac Gly Leu Asp Leu Val Ser Ala Lys Asp Leu Ala Gly Gln Leu Thr Asp 615 620 625 630			2105
cac gac tgg agc ctc ttc aac agt atc cac cag gtg gag ctg atc cac His Asp Trp Ser Leu Phe Asn Ser Ile His Gln Val Glu Leu Ile His 635 640 645			2153
tat gtg ctg ggc ccc cag cat ctg cgg gat gtc acc acc gcc aac ctg Tyr Val Leu Gly Pro Gln His Leu Arg Asp Val Thr Thr Ala Asn Leu 650 655 660			2201
gag cgc ttc atg cgc cgc ttc aat gag ctg cag tac tgg gtg gcc acc Glu Arg Phe Met Arg Arg Phe Asn Glu Leu Gln Tyr Trp Val Ala Thr 665 670 675			2249
gag ctg tgt ctc tgc ccc gtg ccc ggc ccc cgg gcc cag ctg ctc aaa Glu Leu Cys Leu Cys Pro Val Pro Gly Pro Arg Ala Gln Leu Leu Lys 680 685 690			2297
aag ttc att aag ctg gcg gcc cac ctc aag gag cag aag aat gtc aat Lys Phe Ile Lys Leu Ala Ala His Leu Lys Glu Gln Lys Asn Val Asn 695 700 705 710			2345
tcc ttc ttt gcc gtc atg ttt ggc ctc agc aac tcg ccc atc agc cgc Ser Phe Phe Ala Val Met Phe Gly Leu Ser Asn Ser Pro Ile Ser Arg 715 720 725			2393

1625 1673 1721 1769 1817 1865 1913 1961 2009 2057 2105 2153 2201 2249 2297 2345 2393

cta gcc cac acc tgg gag cgg ctg cct cac aaa gtc cgg aag ctg tac 2441
 Leu Ala His Thr Trp Glu Arg Leu Pro His Lys Val Arg Lys Leu Tyr
 730 735 740

tcc gcc ctc gag agg ctg ctg gat ccc tca tgg aac cac cgg gta tac 2489
 Ser Ala Leu Glu Arg Leu Leu Asp Pro Ser Trp Asn His Arg Val Tyr
 745 750 755

cga ctg gcc ctc gcc aag ctc tcc cct cct gtc atc ccc ttc atg ccc 2537
 Arg Leu Ala Leu Ala Lys Leu Ser Pro Pro Val Ile Pro Phe Met Pro
 760 765 770

ctt ctt ctc aaa gac atg acc ttc att cat gag gga aac cac aca cta 2585
 Leu Leu Leu Lys Asp Met Thr Phe Ile His Glu Gly Asn His Thr Leu
 775 780 785 790

gtg gag aat ctc atc aac ttt gag aag atg aga atg atg gcc aga gcc 2633
 Val Glu Asn Leu Ile Asn Phe Glu Lys Met Arg Met Met Ala Arg Ala
 795 800 805

gcg cgg atg ctg cac cac tgc cga agc cac aac cct gtg cct ctc tca 2681
 Ala Arg Met Leu His His Cys Arg Ser His Asn Pro Val Pro Leu Ser
 810 815 820

cca ctc aga agc cga gtt tcc cac ctc cac gag gac agc cag gtg gcg 2729
 Pro Leu Arg Ser Arg Val Ser His Leu His Glu Asp Ser Gln Val Ala
 825 830 835

agg att tcc aca tgc tgc gag cag tcc ctg agc acc cgg agt cca gcc 2777
 Arg Ile Ser Thr Cys Ser Glu Gln Ser Leu Ser Thr Arg Ser Pro Ala
 840 845 850

agc acc tgg gct tat gtc cag cag ctg aag gtc att gac aac cag cgg 2825
 Ser Thr Trp Ala Tyr Val Gln Gln Leu Lys Val Ile Asp Asn Gln Arg
 855 860 865 870

gaa ctc tcc cgc ctg tcc cga gag ctg gag cca tgaggagggg ctgggactgg 2878
 Glu Leu Ser Arg Leu Ser Arg Glu Leu Glu Pro
 875 880

agctggagca ggcacttgca gccgggaaag ccagggtgtg ccggggccaag atactcacag 2938

gctggccaca gctgggcaag gctctccgtg gagtggactc gagtccctgg agcaggcagt 2998

gtggaggcag ccatcccctg tgatgactgg cagctaagga ggacctcgga gtggacaaaa 3058

gccaggaata acgaatgacc aaggggccaag gaaggaggga cagagaggcc ccaggagtgg 3118

gtggagagtg gagtgcgctg ggacgttgtg tgcaatagag aggtctccac accagatgtc 3178

ttccagattc tgtgcctctg gctttgttgt ccagccaggc ctgcagttta ttttcacagt 3238

ggacagagag agagagagag gctgcatgtg tgtaccgtgt gtggcaaggg cagggccttg 3298

gcctggggca gggggccctg cttttcttcc acagctttct tccaacagca ggcagtgggg 3358

ctgcgggcct gaaaaaaaaa aaaaaaaaaa aaaaaa 3394

<210> 12
 <211> 881
 <212> PRT
 <213> Homo sapiens

<400> 12
 Met Val Leu Arg Arg Met His Arg Pro Arg Ser Cys Ser Tyr Gln Leu
 1 5 10 15

Leu Leu Glu His Gln His Pro Ser Cys Ile Gln Gly Leu Arg Trp Thr
 20 25 30
 Pro Leu Thr Asn Ser Glu Glu Ser Leu Asp Phe Ser Glu Ser Leu Glu
 35 40 45
 Gln Ala Ser Thr Glu Arg Val Leu Arg Ala Gly Arg Gln Leu His Gln
 50 55 60
 His Leu Leu Ala Thr Cys Pro Asn Leu Ile Arg Asp Arg Lys Tyr His
 65 70 75 80
 Leu Arg Leu Tyr Arg Gln Cys Cys Ser Gly Arg Glu Leu Val Asp Gly
 85 90 95
 Ile Leu Ala Leu Gly Leu Gly Val His Ser Arg Ser Gln Val Val Gly
 100 105
 Ile Cys Gln Val Leu Leu Asp Glu Gly Ala Leu Cys His Val Lys His
 115 120 125
 Asp Trp Ala Phe Gln Asp Arg Asp Ala Gln Phe Tyr Arg Phe Pro Gly
 130 135 140
 Pro Glu Pro Glu Pro Val Gly Thr His Glu Met Glu Glu Glu Leu Ala
 145 150 155 160
 Glu Ala Val Ala Leu Leu Ser Gln Arg Gly Pro Asp Ala Leu Leu Thr
 165 170 175
 Val Ala Leu Arg Lys Pro Pro Gly Gln Arg Thr Asp Glu Glu Leu Asp
 180 185 190
 Leu Ile Phe Glu Glu Leu Leu His Ile Lys Ala Val Ala His Leu Ser
 195 200 205
 Asn Ser Val Lys Arg Glu Leu Ala Ala Val Leu Leu Phe Glu Pro His
 210 215 220
 Ser Lys Ala Gly Thr Val Leu Phe Ser Gln Gly Asp Lys Gly Thr Ser
 225 230 235 240
 Trp Tyr Ile Ile Trp Lys Gly Ser Val Asn Val Val Thr His Gly Lys
 245 250 255
 Gly Leu Val Thr Thr Leu His Glu Gly Asp Asp Phe Gly Gln Leu Ala
 260 265 270
 Leu Val Asn Asp Ala Pro Arg Ala Ala Thr Ile Ile Leu Arg Glu Tyr
 275 280 285
 Asn Cys His Phe Leu Arg Val Asp Lys Gln Asp Phe Asn Arg Ile Ile
 290 295 300
 Lys Asp Val Glu Ala Lys Thr Met Arg Leu Glu Glu His Gly Lys Val
 305 310 315 320
 Val Leu Val Leu Glu Arg Ala Ser Gln Gly Ala Gly Pro Ser Arg Pro
 325 330 335
 Pro Thr Pro Gly Arg Asn Arg Tyr Thr Val Met Ser Gly Thr Pro Asp
 340 345 350
 Lys Ile Leu Glu Leu Leu Leu Glu Ala Met Gly Leu Asp Ser Ser Ala
 355 360 365
 His Asp Pro Lys Glu Thr Phe Leu Ser Asp Phe Leu Leu Thr His Arg
 370 375 380

Val Phe Met Pro Ser Ala Gln Leu Cys Ala Ala Leu Leu His His Phe
385 390 395 400

His Val Glu Pro Ala Gly Gly Ser Glu Gln Glu Arg Ser Thr Tyr Val
405 410 415

Cys Asn Lys Arg Gln Gln Ile Leu Arg Leu Val Ser Gln Trp Val Ala
420 425 430

Leu Tyr Gly Ser Met Leu His Thr Asp Pro Val Ala Thr Ser Phe Leu
435 440 445

His Lys Leu Ser Asp Leu Val Gly Arg Asp Thr Arg Leu Ser Asn Leu
450 455 460

Leu Arg Glu Gln Trp Pro Glu Arg Arg Arg Cys His Arg Leu Glu Asn
465 470 475 480

Gly Cys Gly Asn Ala Ser Pro Gln Met Lys Ala Arg Asn Leu Pro Val
485 490 495

Trp Leu Pro Asn Gln Asp Glu Pro Leu Pro Gly Ser Ser Cys Ala Ile
500 505 510

Gln Val Gly Asp Lys Val Pro Tyr Asp Ile Cys Arg Pro Asp His Ser
515 520 525

Val Leu Thr Leu Gln Leu Pro Val Thr Ala Ser Val Arg Glu Val Met
530 535 540

Ala Ala Leu Ala Gln Glu Asp Gly Trp Thr Lys Gly Gln Val Leu Val
545 550 555 560

Lys Val Asn Ser Ala Gly Asp Ala Ile Gly Leu Gln Pro Asp Ala Arg
565 570 575

Gly Val Ala Thr Ser Leu Gly Leu Asn Glu Arg Leu Phe Val Val Asn
580 585 590

Pro Gln Glu Val His Glu Leu Ile Pro His Pro Asp Gln Leu Gly Pro
595 600 605

Thr Val Gly Ser Ala Glu Gly Leu Asp Leu Val Ser Ala Lys Asp Leu
610 615 620

Ala Gly Gln Leu Thr Asp His Asp Trp Ser Leu Phe Asn Ser Ile His
625 630 635 640

Gln Val Glu Leu Ile His Tyr Val Leu Gly Pro Gln His Leu Arg Asp
645 650 655

Val Thr Thr Ala Asn Leu Glu Arg Phe Met Arg Arg Phe Asn Glu Leu
660 665 670

Gln Tyr Trp Val Ala Thr Glu Leu Cys Leu Cys Pro Val Pro Gly Pro
675 680 685

Arg Ala Gln Leu Leu Lys Lys Phe Ile Lys Leu Ala Ala His Leu Lys
690 695 700

Glu Gln Lys Asn Val Asn Ser Phe Phe Ala Val Met Phe Gly Leu Ser
705 710 715 720

Asn Ser Pro Ile Ser Arg Leu Ala His Thr Trp Glu Arg Leu Pro His
725 730 735

Lys Val Arg Lys Leu Tyr Ser Ala Leu Glu Arg Leu Leu Asp Pro Ser
740 745 750

cca aac ctc atc cga gac cgg aag tac cac ctt agg ctc tat cgg cag 473
Pro Asn Leu Ile Arg Asp Arg Lys Tyr His Leu Arg Leu Tyr Arg Gln
75 80 85

tgc Cys	tgc Cys	tct Ser	ggc Gly 90	cgg Arg	gag Glu	ctg Leu	gtg Val	gat Asp 95	ggg Gly	atc Ile	ttg Leu	gcc Ala 100	ctg Leu	gga Gly	ctt Leu	521
ggg Gly	gtc Val	cat His 105	tcc Ser	cgg Arg	agc Ser	caa Gln	gtt Val 110	gtg Val	gga Gly	atc Ile	tgc Cys	cag Gln 115	gtg Val	ctg Leu	ctg Leu	569
gat Asp	gaa Glu 120	ggt Gly	gcc Ala	ctc Leu	tgc Cys	cat His 125	gtg Val	aaa Lys	cac His	gac Asp	tgg Trp 130	gcc Ala	ttc Phe	cag Gln	gac Asp	617
cga Arg 135	gat Asp	gcc Ala	caa Gln	ttc Phe	tac Tyr 140	cgg Arg	ttc Phe	ccc Pro	ggg Gly	ccc Pro 145	gag Glu	ccc Pro	gag Glu	ccc Pro	gtg Val 150	665
gga Gly	act Thr	cat His	gag Glu	atg Met 155	gag Glu	gag Glu	gag Glu	ttg Leu	gcc Ala 160	gaa Glu	gct Ala	gtg Val	gcc Ala	ctg Leu 165	ctc Leu	713
tcc Ser	cag Gln	cgg Arg	ggg Gly 170	cct Pro	gac Asp	gcc Ala	ctg Leu	ctc Leu 175	act Thr	gtg Val	gca Ala	ctt Leu	cga Arg 180	aag Lys	ccc Pro	761
cca Pro	ggt Gly	cag Gln 185	cgc Arg	acg Thr	gat Asp	gaa Glu	gag Glu 190	ctg Leu	gac Asp	ctc Leu	atc Ile	ttt Phe 195	gag Glu	gag Glu	ctg Leu	809
ctg Leu	cac His 200	atc Ile	aag Lys	gct Ala	gtg Val	gcc Ala 205	cac His	ctc Leu	tcc Ser	aac Asn	tgc Ser 210	gtg Val	aag Lys	cga Arg	gaa Glu	857
tta Leu 215	gcg Ala	gct Ala	gtt Val	ctg Leu 220	ctc Leu	ttt Phe	gaa Glu	cca Pro	cac His	agc Ser 225	aag Lys	gca Ala	ggg Gly	acc Thr	gtg Val 230	905
ttg Leu	ttc Phe	agc Ser	cag Gln 235	ggg Gly 235	gac Asp	aag Lys	ggc Gly	act Thr 240	tgc Ser	tgg Trp	tac Tyr	att Ile	atc Ile	tgg Trp 245	aag Lys	953
gga Gly	tct Ser	gtc Val	aac Asn 250	gtg Val	gtg Val	acc Thr	cat His	ggc Gly 255	aag Lys	ggg Gly	ctg Leu	gtg Val	acc Thr 260	acc Thr	ctg Leu	1001
cat His	gag Glu	gga Gly 265	gat Asp	gat Asp	ttt Phe	gga Gly	cag Gln 270	ctg Leu	gct Ala	ctg Leu	gtg Val	aat Asn 275	gat Asp	gca Ala	ccc Pro	1049
cgg Arg	gca Ala 280	gcc Ala	acc Thr	atc Ile	atc Ile	ctg Leu 285	cga Arg	gaa Glu	tac Tyr	aac Asn	tgt Cys 290	cat His	ttc Phe	ctg Leu	cgt Arg	1097
gtg Val 295	gac Asp	aag Lys	cag Gln	gac Asp	ttc Phe 300	aac Asn	cgt Arg	atc Ile	atc Ile	aag Lys 305	gat Asp	gtg Val	gag Glu	gca Ala	aag Lys 310	1145
acc Thr	atg Met	cgg Arg	ctg Leu	gaa Glu 315	gaa Glu	cat His	ggc Gly	aaa Lys	gtg Val 320	gtg Val	ctg Leu	gtg Val	ctg Leu	gag Glu 325	aga Arg	1193
gcc Ala	tct Ser	cag Gln 330	ggc Gly	gcc Ala	ggc Gly	cct Pro	tcc Ser	cga Arg 335	ccc Pro	cca Pro	acc Thr	cca Pro	ggc Gly 340	agg Arg	aac Asn	1241
cgg Arg	tat Tyr	aca Thr 345	gtg Val	atg Met	tct Ser	ggc Gly	act Thr 350	cca Pro	gat Asp	aag Lys	atc Ile	cta Leu 355	gag Glu	ctt Leu	ctg Leu	1289
ttg Leu	gag Glu	gcc Ala	atg Met	gga Glu	cta Thr	gat Asp	tcc Ser	agt Glu	gct Ala	cat Glu	gac Ala	cca Ala	aaa Glu	gaa Glu	aca Thr	1337

bioRxiv preprint doi: <https://doi.org/10.1101/111111>; this version posted November 11, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Leu 360	Glu 360	Ala 360	Met 360	Gly 360	Leu 365	Asp 365	Ser 370	Ser 370	Ala 370	His 370	Asp 370	Pro 370	Lys 370	Glu 370	Thr 370	
ttc Phe 375	ctc Leu 375	agc Ser 375	gac Asp 375	ttc Phe 380	ctc Leu 380	ctg Leu 380	acc Thr 380	cac His 385	agg Arg 385	gtc Val 385	ttc Phe 390	atg Met 390	ccc Pro 390	agc Ser 390	gcc Ala 390	1385
caa Gln 395	ctc Leu 395	tgc Cys 395	gct Ala 395	gcc Ala 395	ctt Leu 400	ctg Leu 400	cac His 400	cac His 400	ttc Phe 400	cat His 400	gtg Val 405	gag Glu 405	cct Pro 405	gcg Ala 405	ggg Gly 405	1433
ggc Gly 410	agc Ser 410	gag Glu 410	cag Gln 410	gag Glu 415	cgc Arg 415	agc Ser 415	acc Thr 415	tac Tyr 415	gtc Val 420	tgc Cys 420	aac Asn 420	aag Lys 420	agg Arg 420	cag Gln 425	cag Gln 425	1481
atc Ile 425	ttg Leu 425	cgg Arg 425	ctg Leu 430	gtc Val 430	agc Ser 430	cag Gln 430	tgg Trp 430	gtg Val 435	gcc Ala 435	ctg Leu 435	tat Tyr 435	ggc Gly 435	tcc Ser 440	atg Met 440	ctc Leu 440	1529
cac His 440	act Thr 440	gac Asp 445	cct Pro 445	gtg Val 445	gcc Ala 445	acc Thr 445	agc Ser 445	ttc Phe 450	ctc Leu 450	cag Gln 450	aaa Lys 450	ctc Leu 455	tca Ser 455	gac Asp 455	ctg Leu 455	1577
gtg Val 455	ggc Gly 460	agg Arg 460	gac Asp 460	acc Thr 460	cga Arg 460	ctc Leu 465	agc Ser 465	aac Asn 465	ctg Leu 465	ctg Leu 465	agg Arg 465	gag Glu 470	cag Gln 470	tgg Trp 470	cca Pro 470	1625
gag Glu 475	agg Arg 475	cgg Arg 475	cga Arg 475	tgc Cys 475	cac His 475	agg Arg 480	ttg Leu 480	gag Glu 480	aat Asn 480	ggc Gly 480	tgt Cys 485	ggg Gly 485	aat Asn 485	gca Ala 485	tct Ser 485	1673
cct Pro 490	cag Gln 490	atg Met 490	aag Lys 490	gtg Val 495	tct Ser 495	gcc Ala 495	tgg Trp 495	ccc Pro 495	cag Gln 500	ttt Phe 500	ctt Leu 500	tcc Ser 500	tct Ser 500	gct Ala 505	cct Pro 505	1721
cct Pro 505	gga Gly 505	ctg Leu 505	cag Gln 510	gca Ala 510	cct Pro 510	cct Pro 510	tcg Ser 510	ccc Pro 510	cct Pro 510	gac Asp 515	cct Pro 515	gag Glu 515	ggg Gly 515	ctc Leu 520	tgt Cys 520	1769
ggg Gly 520	cgt Arg 520	ggg Gly 520	aag Lys 520	ctc Leu 525	tcc Ser 525	tcc Ser 525	cac His 525	aga Arg 530	cac His 530	acc Thr 530	ctt Leu 530	ggg Gly 530	tct Ser 535	ctg Leu 535	ata Ile 535	1817
ggg Gly 535	gtt Val 535	cac His 540	ggg Gly 540	gcc Ala 540	ctt Leu 540	gct Ala 540	gca Ala 540	tgt Cys 545	ggg Gly 545	gcc Ala 545	ctg Leu 545	ggc Gly 550	cag Gln 550	gcc Ala 550	gtg Val 550	1865
cca Pro 555	gga Gly 555	ggc Gly 555	gca Ala 555	gag Glu 555	gcc Ala 555	taag	gttg	gcc	tcct	ctcg	ccact	ccct				1913
gactcaatgg gccttttatt ccttttg gga ggtaattcat gccacaggt agagcctggg																1973
agatgaggaa tgggtgctg agttggcccc tgaggccac ggggtcctg gtggccagt																2033
ctgtgggagc tcagaggaa gcgaggcccc tcctggaaa gtcagagggg gcccaggcc																2093
attctcagct ggggcttaa ggaagcaaa gaggggatga gaagatgtga ctgcagccag																2153
gatttggttg aggaatggt ggagaaagaa ctggaaagg agattagggt ggtggtggg																2213
agggtgcac tgggaataaa atccaggggc ctcaccatcc agcaggctcc catggtccc																

```

cagacaggcc ccctcccgtt tccctatccg gagcaccctc ccctttctcc caccacgctc 2513
tgccttgggtg tcctctgtga tttttttcaa agcactctcg ccatgtggag ttgttttctt 2573
tgcctacttg ctttctgtcc ttttggtatct ctcccatccc actggagagt tgacttccct 2633
cctgcggggcc cccacctccg cttctccttg ctcacatgct gcctggctct tggtcgacgc 2693
tcagaaactg atgtgggctc accctcgcag gctgggcact gcgctagggtg ctgatgctgc 2753
acctcattca accctggcct gtggtggctg cagtcaggca aggggcagag cagttcaaca 2813
acccgtgatg tgcagtgaca cgggacactt tccctgaatg tgcactggag tcggctcttc 2873
caactgtagg tccgctgcct taaaccagga aggggaagag ccaactccag tgcagattca 2933
gggaaagtgt cccgtgtaag cgactactgg agttgagagc tgcagggggc aagggtggag 2993
gaagaagagg attcctggtg ggcaaaccga ggaggaaagg gaaccagga ggaggatccc 3053
tgaggaggcg agagtaagag taggacccca gagcaggcag gaggtgggca gggccccctg 3113
ggcaaggaca ggggtctgat ctttgtccta agagcagtgc aaaccattaa aggggtttga 3173
gccaaagttg gcaaagtctg atttgtgttt tgaataggct agctaaagag cctgggcttt 3233
ctccagcaga caatggacac cactgaagtc gagtgcggga gcttcacagg gctctggtca 3293
ggcagattgg aggaggagg ctggaggcac ctgggaggga ggaggaggag catagttggg 3353
ccattggagt agtaaggagt gaggcctcgt gggcagggtg aactggagtt agcctgggca 3413
ggggaggagg gggcacagaa gaccacttc aaagaagaat cttcaaagca agatgacaag 3473
ctacggaagt gtggaggagc ctggagctgg ggagaatggc tggggaacag agtggctctg 3533
gaggcaggat gcaaggatcc gatgggtata tggagtgtga ggtaatgggt cattcatgtg 3593
gaaggatgca gggggttttt gagaccaggg tttggaagag agttcagcac tgctggtagt 3653
tttggaatc acccatgtgc aggcgacaca tgaggcagta aggaactctg caggggtccc 3713
tgagatttgg aaatgtaggg aagagcaatg gattgaggtc cgaacctgga ggatctgcta 3773
tacgcagagc tgggaggagg gacagagtca gtaccagagt cggaaaaaag caggggtggga 3833
aggggaacct gagtccaggag acttgcttgg caggcgctgc cctgccagca gaggcctgac 3893
agtggtttcc atgaactgca tccctgctgt gggctgggac agggccactg acacagtatc 3953
ggagcacaga aggggaaagg agcaggaggg attccaactc tgccagttag cagctgtgtg 4013
gctttgggca tgttacttaa cctctctgag cctcatttat ttcattcata aaatggaaat 4073
aaaaataata cttttgtcaa aaaaaaaaaa aaaaaa 4109

```

<210> 14
 <211> 556
 <212> PRT
 <213> Homo sapiens

<400> 14
 Met Val Leu Arg Arg Met His Arg Pro Arg Ser Cys Ser Tyr Gln Leu
 1 5 10 15
 Leu Leu Glu His Gln His Pro Ser Cys Ile Gln Gly Leu Arg Trp Thr
 20 25 30

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

Pro Leu Thr Asn Ser Glu Glu Ser Leu Asp Phe Ser Glu Ser Leu Glu
35 40 45

Gln Ala Ser Thr Glu Arg Val Leu Arg Ala Gly Arg Gln Leu His Gln
50 55 60

His Leu Leu Ala Thr Cys Pro Asn Leu Ile Arg Asp Arg Lys Tyr His
65 70 75 80

Leu Arg Leu Tyr Arg Gln Cys Cys Ser Gly Arg Glu Leu Val Asp Gly
85 90 95

Ile Leu Ala Leu Gly Leu Gly Val His Ser Arg Ser Gln Val Val Gly
100 105 110

Ile Cys Gln Val Leu Leu Asp Glu Gly Ala Leu Cys His Val Lys His
115 120 125

Asp Trp Ala Phe Gln Asp Arg Asp Ala Gln Phe Tyr Arg Phe Pro Gly
130 135 140

Pro Glu Pro Glu Pro Val Gly Thr His Glu Met Glu Glu Glu Leu Ala
145 150 155 160

Glu Ala Val Ala Leu Leu Ser Gln Arg Gly Pro Asp Ala Leu Leu Thr
165 170 175

Val Ala Leu Arg Lys Pro Pro Gly Gln Arg Thr Asp Glu Glu Leu Asp
180 185 190

Leu Ile Phe Glu Glu Leu Leu His Ile Lys Ala Val Ala His Leu Ser
195 200 205

Asn Ser Val Lys Arg Glu Leu Ala Ala Val Leu Leu Phe Glu Pro His
210 215 220

Ser Lys Ala Gly Thr Val Leu Phe Ser Gln Gly Asp Lys Gly Thr Ser
225 230 235 240

Trp Tyr Ile Ile Trp Lys Gly Ser Val Asn Val Val Thr His Gly Lys
245 250 255

Gly Leu Val Thr Thr Leu His Glu Gly Asp Asp Phe Gly Gln Leu Ala
260 265 270

Leu Val Asn Asp Ala Pro Arg Ala Ala Thr Ile Ile Leu Arg Glu Tyr
275 280 285

Asn Cys His Phe Leu Arg Val Asp Lys Gln Asp Phe Asn Arg Ile Ile
290 295 300

Lys Asp Val Glu Ala Lys Thr Met Arg Leu Glu Glu His Gly Lys Val
305 310 315 320

Val Leu Val Leu Glu Arg Ala Ser Gln Gly Ala Gly Pro Ser Arg Pro
325 330 335

Pro Thr Pro Gly Arg Asn Arg Tyr Thr Val Met Ser Gly Thr Pro Asp
340 345 350

Lys Ile Leu Glu Leu Leu Leu Glu Ala Met Gly Leu Asp Ser Ser Ala
355 360 365

His Asp Pro Lys Glu Thr Phe Leu Ser Asp Phe Leu Leu Thr His Arg
370 375 380

Val Phe Met Pro Ser Ala Gln Leu Cys Ala Ala Leu Leu His His Phe
385 390 395 400

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

His Val Glu Pro Ala Gly Gly Ser Glu Gln Glu Arg Ser Thr Tyr Val
405 410 415
Cys Asn Lys Arg Gln Gln Ile Leu Arg Leu Val Ser Gln Trp Val Ala
420 425 430
Leu Tyr Gly Ser Met Leu His Thr Asp Pro Val Ala Thr Ser Phe Leu
435 440 445
Gln Lys Leu Ser Asp Leu Val Gly Arg Asp Thr Arg Leu Ser Asn Leu
450 455 460
Leu Arg Glu Gln Trp Pro Glu Arg Arg Arg Cys His Arg Leu Glu Asn
465 470 475 480
Gly Cys Gly Asn Ala Ser Pro Gln Met Lys Val Ser Ala Trp Pro Gln
485 490 495
Phe Leu Ser Ser Ala Pro Pro Gly Leu Gln Ala Pro Pro Ser Pro Pro
500 505 510
Asp Pro Glu Gly Leu Cys Gly Arg Gly Lys Leu Ser Ser His Arg His
515 520 525
Thr Leu Gly Ser Leu Ile Gly Val His Gly Ala Leu Ala Ala Cys Gly
530 535 540
Ala Leu Gly Gln Ala Val Pro Gly Gly Ala Glu Ala
545 550 555

<210> 15
<211> 1966
<212> DNA
<213> Rattus norvegicus

<220>
<221> CDS
<222> (3)..(875)
<223> cAMP-GEFII

<400> 15
aa ggt gtg ctc aaa cct aat gat gtt tca gta ttt acg acg ctc acc 47
Gly Val Leu Lys Pro Asn Asp Val Ser Val Phe Thr Thr Leu Thr 15
1 5 10
att aat gga cgc ctg ttt gcc tgc ccg cga gag caa ttc gac tca ctg 95
Ile Asn Gly Arg Leu Phe Ala Cys Pro Arg Glu Gln Phe Asp Ser Leu 20 25 30
act ccc ttg cca gaa cag gag ggc ccg acc act ggg aca gtg ggg acg 143
Thr Pro Leu Pro Glu Gln Glu Gly Pro Thr Thr Gly Thr Val Gly Thr 35 40 45
ttt gaa ctg atg agc tcg aaa gac ttg gcg tac cag atg aca acg tat 191
Phe Glu Leu Met Ser Ser Lys Asp Leu Ala Tyr Gln Met Thr Thr Tyr 50 55 60
gac tgg gaa ctc ttc aac tgt gtg ctc gag ctg gag cta atc tac cac 239
Asp Trp Glu Leu Phe Asn Cys Val Leu Glu Leu Glu Leu Ile Tyr His 65 70 75
aca ttt gga agg cat aat ttt aaa aag acc aca gca aac ttg gat ttg 287
Thr Phe Gly Arg His Asn Phe Lys Lys Thr Thr Ala Asn Leu Asp Leu 80 85 90 95
ttc ctg agg aga ttt aat gaa att cag ttt tgg gtt gtc act gag atc 335
Phe Leu Arg Arg Phe Asn Glu Ile Gln Phe Trp Val Val Thr Glu Ile 100 105 110

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tgc ctt tgt tcc cag ctc agc aag cgt gtt cag ctt ttg aaa aaa tgt	383
Cys Leu Cys Ser Gln Leu Ser Lys Arg Val Gln Leu Leu Lys Lys Cys	
115 120 125	
atc aag ata gcg gct cac tgc aag gag tac aaa aac ttg aat tcc ttc	431
Ile Lys Ile Ala Ala His Cys Lys Glu Tyr Lys Asn Leu Asn Ser Phe	
130 135 140	
ttc ggc atc gtc atg ggg ctc agt aac gtg gct gag agc cgc ctg gca	479
Phe Gly Ile Val Met Gly Leu Ser Asn Val Ala Glu Ser Arg Leu Ala	
145 150 155	
tta aca tgg gag aaa ctg ccg agc aag ttt aag aag ttc tat gcg gag	527
Leu Thr Trp Glu Lys Leu Pro Ser Lys Phe Lys Lys Phe Tyr Ala Glu	
160 165 170 175	
ttt gag agc tta atg gat cct tcc aga aat cac aag gcg tac agg ctg	575
Phe Glu Ser Leu Met Asp Pro Ser Arg Asn His Lys Ala Tyr Arg Leu	
180 185 190	
aca gca gct aaa ctg gag ccc ccc ctc atc cct ttc atg ccc ttg ctt	623
Thr Ala Ala Lys Leu Glu Pro Pro Leu Ile Pro Phe Met Pro Leu Leu	
195 200 205	
att aaa gat atg aca ttt act cat gag ggg aac aag aca ttc att gac	671
Ile Lys Asp Met Thr Phe Thr His Glu Gly Asn Lys Thr Phe Ile Asp	
210 215 220	
aat cta gta aac ttt gaa aaa atg cgc atg att gca aat act gcc aga	719
Asn Leu Val Asn Phe Glu Lys Met Arg Met Ile Ala Asn Thr Ala Arg	
225 230 235	
acg gtg cgc tac tac agg agc cag cca ttc aat ccg gat gct gct caa	767
Thr Val Arg Tyr Tyr Arg Ser Gln Pro Phe Asn Pro Asp Ala Ala Gln	
240 245 250 255	
gct aat aag aac cat cag gat gtc cgg agt tat gta cgg caa tta aat	815
Ala Asn Lys Asn His Gln Asp Val Arg Ser Tyr Val Arg Gln Leu Asn	
260 265 270	
gtg att gac aac cag aga act tta tca cag atg tca cac aga tta gag	863
Val Ile Asp Asn Gln Arg Thr Leu Ser Gln Met Ser His Arg Leu Glu	
275 280 285	
cct cgc agg cca tagacatctg cagtgccag agtgatgctc cgtctccagt	915
Pro Arg Arg Pro	
290	

ccacaatctt tcaaaagatg ctgtgtatgc tactactgac tgtgttgcta ctagagaatt 975

ccccagaatg agcaagagac acctcctgag agcccctcgg ggccacatcc tgctttccga 1035

ccacacagga gaaggatctg tcttgtctaa cgcgacatg ccgtagctta ggaaccatca 1095

gctgtagtca tcttcttcac gggtggcaca ccaccgcagg ctcacgtgaa ggcataacct 1155

ggcgaggcta caccaggccc ctgacatccc ttcccaggct gttgcagcat gagactgtcc 1215

cgtggatagg tttgacttgg aatcgctgca atgatataat tgaatgattt gtttacttag 1275

caccttattt ggggtctggg ttctggggag ggtgttgacc ataaaagtcc aaattatcca 1335

tcatgttctt ccacgtcat aatcttacct ctgaaggaat ggaacctcat cacaacacta 1395

tgaacatta tgactgttca gtctgtaatt tcggaatgat ctatagaata atatgtttac 1455

attgtaactt ttaaaaactt acaaatcagg attacacaca tgagaattcc actaagaaac 1515

```

accaagggttc ttaatatcgc cagcggttaag atagaaagta acatttcaga agagcacaat 1575
atacacaaaaa catttttttca aattgaaata ttttcctggg cattaaaaaa cttttccact 1635
acaaattttat tggtactgat gaaaaaaaaa gcatattttc tggacttaaa tgttattaca 1695
aaaatcttaa ttttcagcaa ttgttttgca ctttcagata gattgtaaat aggttatgca 1755
gtcaatggta tagaattatt tatttgctac ataatagaca ttgtgccaaa taattccttt 1815
ttattttattt tattcagtat gaaattttgg agtacatttt ttctgttttc ttaattagac 1875
tacatttaat gtataggaat tgtatgtaca tatctcttct gtaaataaca gccagtatct 1935
tcattaaata tacttgacaa gaaaaaaaaa a 1966

```

```

<210> 16
<211> 291
<212> PRT
<213> Rattus norvegicus

```

```

<400> 16
Gly Val Leu Lys Pro Asn Asp Val Ser Val Phe Thr Thr Leu Thr Ile
  1          5          10          15
Asn Gly Arg Leu Phe Ala Cys Pro Arg Glu Gln Phe Asp Ser Leu Thr
  20          25          30
Pro Leu Pro Glu Gln Glu Gly Pro Thr Thr Gly Thr Val Gly Thr Phe
  35          40          45
Glu Leu Met Ser Ser Lys Asp Leu Ala Tyr Gln Met Thr Thr Tyr Asp
  50          55          60
Trp Glu Leu Phe Asn Cys Val Leu Glu Leu Glu Leu Ile Tyr His Thr
  65          70          75          80
Phe Gly Arg His Asn Phe Lys Lys Thr Thr Ala Asn Leu Asp Leu Phe
  85          90          95
Leu Arg Arg Phe Asn Glu Ile Gln Phe Trp Val Val Thr Glu Ile Cys
 100          105          110
Leu Cys Ser Gln Leu Ser Lys Arg Val Gln Leu Leu Lys Lys Cys Ile
 115          120          125
Lys Ile Ala Ala His Cys Lys Glu Tyr Lys Asn Leu Asn Ser Phe Phe
 130          135          140
Gly Ile Val Met Gly Leu Ser Asn Val Ala Glu Ser Arg Leu Ala Leu
 145          150          155          160
Thr Trp Glu Lys Leu Pro Ser Lys Phe Lys Lys Phe Tyr Ala Glu Phe
 165          170          175
Glu Ser Leu Met Asp Pro Ser Arg Asn His Lys Ala Tyr Arg Leu Thr
 180          185          190
Ala Ala Lys Leu Glu Pro Pro Leu Ile Pro Phe Met Pro Leu Leu Ile
 195          200          205
Lys Asp Met Thr Phe Thr His Glu Gly Asn Lys Thr Phe Ile Asp Asn
 210          215          220
Leu Val Asn Phe Glu Lys Met Arg Met Ile Ala Asn Thr Ala Arg Thr
 225          230          235          240
Val Arg Tyr Tyr Arg Ser Gln Pro Phe Asn Pro Asp Ala Ala Gln Ala

```

123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100101102103104105106107108109110111112113114115116117118119120121122123124125126127128129130131132133134135136137138139140141142143144145146147148149150151152153154155156157158159160161162163164165166167168169170171172173174175176177178179180181182183184185186187188189190191192193194195196197198199200201202203204205206207208209210211212213214215216217218219220221222223224225226227228229230231232233234235236237238239240241242243244245246247248249250251252253254255256257258259260261262263264265266267268269270271272273274275276277278279280281282283284285286287288289290291292293294295296297298299300

	245		250		255
Asn Lys Asn His Gln Asp Val Arg Ser Tyr Val Arg Gln Leu Asn Val	260	265	270		
Ile Asp Asn Gln Arg Thr Leu Ser Gln Met Ser His Arg Leu Glu Pro	275	280	285		
Arg Arg Pro	290				

<210> 17
 <211> 3013
 <212> DNA
 <213> Homo sapiens

<220>
 <221> CDS
 <222> (407)..(2953)
 <223> cAMP-GEFII

<400> 17
 gatccagcga agatgtggat ataattcttca ctgcactgaa agaagttaaa gcttttgaga 60
 aatttcacccc aaatctcctt catcagatgtt gcttatgtgg ttattatgag aatctggaaa 120
 aggggaataac attatttcgc cagggtgata tggaacaaac tggatatgctg cctggcaggg 180
 tctttggatg ttaaagtatc tgagaccagc agtcaccagg atgctgtgac catctgtacc 240
 ctgggaattg ggacggcctt tggagagtcc attctggaca acacaccccg ccatgcaacc 300
 atcgttacca gggagagcag tgaactgctc cgcacgcagc agaaggactt caaggcacta 360
 tgggagaaat atcgacagta tatggcagga cttctggctc ctcctt atg gta tta 415
 Met Val Leu
 1
 tgg aaa cgg gct cta aca atg aca gga ttc ctg aca agg aga aca cac 463
 Trp Lys 5 Arg Ala Leu Thr Met 10 Thr Gly Phe Leu Thr 15 Arg Arg Thr His
 ctc att gaa cct cac gtt cct ctt cgt cct gct aac acc att acc aag 511
 Leu Ile Glu Pro His Val 25 Pro Leu Arg Pro Ala 30 Asn Thr Ile Thr Lys 35
 20
 gtc cct tca gag aag atc ctc aga gct gga aaa att tta cga aat gcc 559
 Val Pro Ser Glu Lys 40 Ile Leu Arg Ala 45 Gly Lys Ile Leu Arg Asn Ala 50
 att ctc tct cga gca cct cac atg ata aga gat aga aaa tac cac cta 607
 Ile Leu Ser Arg 55 Ala Pro His Met Ile 60 Arg Asp Arg Lys Tyr 65 His Leu
 aag aca tac aga caa tgc tgt gtg gga act gaa ctg gtg gac tgg atg 655
 Lys Thr Tyr 70 Arg Gln Cys Cys Val 75 Gly Thr Glu Leu Val 80 Asp Trp Met
 atc gac gag aca cca tgt gtt cac tcc cgg act caa gct gtt ggc atg 703
 Ile Asp 85 Glu Thr Pro Cys Val 90 His Ser Arg Thr Gln 95 Ala Val Gly Met
 tgg caa gtc ctg tta gaa gat ggt gtt ctc aac cac gtg gac cag gag 751
 Trp Gln Val Leu Leu 105 Glu Asp Gly Val Leu 110 Asn His Val Asp Gln Glu 115
 cac cat ttc caa gac ttt tat tta ttc tat cga ttt ctg gat gat gag 799
 His His Phe Gln Asp Phe Tyr Leu Phe Tyr Arg Phe Leu Asp Asp Glu

120								125						130					
cac His	gag Glu	gat Asp	gcc Ala 135	cct Pro	ttg Leu	cct Pro	act Thr	gag Glu 140	gag Glu	gag Glu	aag Lys	aag Lys	gag Glu 145	tgt Cys	gat Asp	847			
gag Glu	gag Glu	ctc Leu 150	cag Gln	gac Asp	acc Thr	atg Met	ctg Leu 155	ctg Leu	ctg Leu	tca Ser	cag Gln	atg Met 160	ggc Gly	ccc Pro	gac Asp	895			
gcc Ala	cac His 165	atg Met	agg Arg	atg Met	atc Ile	ctt Leu 170	cgc Arg	aaa Lys	cca Pro	cct Pro	ggc Gly 175	cag Gln	agg Arg	act Thr	gtg Val	943			
gat Asp 180	gac Asp	cta Leu	gag Glu	att Ile	atc Ile 185	tat Tyr	gag Glu	gag Glu	ctt Leu	ctt Leu 190	cat His	att Ile	aaa Lys	gcc Ala	tta Leu 195	991			
tcc Ser	cat His	ctt Leu	tct Ser	acc Thr 200	aca Thr	gtg Val	aaa Lys	cga Arg	gag Glu 205	tta Leu	gca Ala	ggg Gly	ggt Val	ctc Leu 210	att Ile	1039			
ttt Phe	gag Glu	tct Ser	cac His 215	gcc Ala	aaa Lys	gga Gly	ggg Gly	act Thr 220	gtg Val	ttg Leu	ttt Phe	aac Asn	cag Gln 225	ggg Gly	gaa Glu	1087			
gaa Glu	ggg Gly	acc Thr 230	tcc Ser	tgg Trp	tac Tyr	att Ile	att Ile 235	cta Leu	aaa Lys	gga Gly	tca Ser	gtg Val 240	aat Asn	gta Val	gtc Val	1135			
att Ile	tac Tyr 245	ggc Gly	aag Lys	ggg Gly	gtg Val	gtc Val 250	tgc Cys	acc Thr	ctg Leu	cat His	gaa Glu 255	gga Gly	gat Asp	gac Asp	ttc Phe	1183			
ggc Gly 260	aag Lys	tta Leu	gca Ala	cta Leu	gtg Val 265	aat Asn	gat Asp	gcc Ala	cca Pro	cga Arg 270	gct Ala	gcc Ala	tct Ser	atc Ile	gtc Val 275	1231			
tta Leu	cga Arg	gaa Glu	gat Asp	aac Asn 280	tgc Cys	cat His	ttc Phe	tta Leu	aga Arg 285	gta Val	gac Asp	aag Lys	gag Glu	gat Asp 290	ttc Phe	1279			
aac Asn	cgg Arg	atc Ile	cta Leu 295	agg Arg	gac Asp	gtg Val	gag Glu	gcg Ala 300	aat Asn	aca Thr	gtc Val	aga Arg	ctt Leu 305	aaa Lys	gaa Glu	1327			
cat His	gac Asp	caa Gln 310	gat Asp	gtc Val	ttg Leu	gtg Val	ctg Leu 315	gag Glu	aag Lys	gtc Val	cca Pro	gca Ala 320	ggg Gly	aac Asn	aga Arg	1375			
gct Ala	tct Ser 325	aat Asn	caa Gln	gga Gly	aac Asn	tca Ser 330	cag Gln	cct Pro	cag Gln	caa Gln	aag Lys 335	tat Tyr	act Thr	gtg Val	atg Met	1423			
tca Ser 340	gga Gly	aca Thr	cct Pro	gaa Glu	aaa Lys 345	att Ile	tta Leu	gag Glu	cat His	ttt Phe 350	cta Leu	gaa Glu	aca Thr	ata Ile	cgc Arg 355	1471			
ctt Leu	gag Glu	gca Ala	act Thr	tta Leu 360	aat Asn	gaa Glu	gca Ala	aca Thr	gat Asp 365	tct Ser	ggt Val	tta Leu	aat Asn	gac Asp 370	ttt Phe	1519			
att Ile	atg Met	atg Met	cac His 375	tgt Cys	gtt Val	ttt Phe	atg Met	cca Pro 380	aat Asn	acc Thr	cag Gln	ctt Leu	tgc Cys 385	ccg Pro	gca Ala	1567			
ctg Leu	gtg Val	gcc Ala 390	cac His	tac Tyr	cac His	gca Ala	cag Gln 395	cct Pro	tca Ser	caa Gln	ggg Gly	aca Thr 400	gaa Glu	cag Gln	gag Glu	1615			

120 125 130 847 895 943 991 1039 1087 1135 1183 1231 1279 1327 1375 1423 1471 1519 1567 1615

aaa Lys	atg Met 405	gat Asp	tat Tyr	gcc Ala	ctc Leu	aac Asn 410	aat Asn	aag Lys	agg Arg	cga Arg	gtc Val 415	atc Ile	cgc Arg	ctg Leu	gtt Val	1663
cta Leu 420	cag Gln	tgg Trp	gct Ala	gcc Ala	atg Met 425	tat Tyr	gga Gly	gac Asp	ctc Leu	ctg Leu 430	caa Gln	gag Glu	gat Asp	gac Asp	gta Val 435	1711
tct Ser	atg Met	gcc Ala	ttc Phe	ctg Leu 440	gag Glu	gag Glu	ttt Phe	tat Tyr	gta Val 445	tct Ser	gta Val	tca Ser	gat Asp	gat Asp 450	gcc Ala	1759
cgg Arg	atg Met	att Ile	gct Ala 455	gcc Ala	ctc Leu	aag Lys	gag Glu	caa Gln 460	ctg Leu	cca Pro	gag Glu	ttg Leu	gag Glu 465	aag Lys	att Ile	1807
gtc Val	aag Lys	caa Gln 470	atc Ile	tca Ser	gaa Glu	gat Asp	gca Ala 475	aag Lys	gca Ala	cca Pro	caa Gln	aag Lys 480	aag Lys	cac His	aag Lys	1855
gtt Val	ctt Leu 485	ttg Leu	caa Gln	cag Gln	ttc Phe	aat Asn 490	acg Thr	ggc Gly	gat Asp	gag Glu	aga Arg 495	gcc Ala	cag Gln	aag Lys	cgc Arg	1903
cag Gln 500	cct Pro	atc Ile	cgc Arg	ggc Gly	tct Ser 505	gat Asp	gaa Glu	gtt Val	ctg Leu	ttt Phe 510	aag Lys	gtc Val	tat Tyr	tgc Cys	atg Met 515	1951
gac Asp	cac His	acc Thr	tac Tyr	aca Thr 520	acc Thr	att Ile	cgg Arg	gtg Val	cca Pro 525	gtg Val	gcc Ala	act Thr	tcg Ser	gtg Val 530	aag Lys	1999
gaa Glu	gtc Val	atc Ile	agt Ser 535	gca Ala	gtt Val	gcc Ala	gac Asp	aag Lys 540	ctg Leu	ggc Gly	tcc Ser	ggg Gly	gag Glu 545	ggc Gly	ctg Leu	2047
atc Ile	ata Ile	gtc Val 550	aag Lys	atg Met	agt Ser	tcc Ser	gga Gly 555	gga Gly	gaa Glu	aag Lys	gtg Val 560	gtg Val	ctc Leu	aaa Lys	cct Pro	2095
aat Asn 565	gat Asp	gtt Val	tca Ser	gta Val	ttt Phe	acg Thr 570	acg Thr	ctc Leu	acc Thr	att Ile	aat Asn 575	gga Gly	cgc Arg	ctg Leu	ttt Phe	2143
gct Ala 580	tgc Cys	ccg Pro	cga Arg	gag Glu	caa Gln 585	ttc Phe	gat Asp	tca Ser	ctg Leu	act Thr 590	ccc Pro	tta Leu	cca Pro	gaa Glu	cag Gln 595	2191
gaa Glu	ggc Gly	cca Pro	act Thr	gtt Val 600	gga Gly	aca Thr	gtg Val	gga Gly	act Thr 605	ttt Phe	gaa Glu	ctg Leu	atg Met	agc Ser 610	tcc Ser	2239
aaa Lys	gat Asp	tta Leu	gca Ala 615	tac Tyr	cag Gln	atg Met	aca Thr	att Ile 620	tat Tyr	gat Asp	tgg Trp	gaa Glu	ctc Leu 625	ttc Phe	aac Asn	2287
tgc Cys	gtg Val	cat His 630	gag Glu	ctg Leu	gag Glu	cta Leu	atc Ile 635	tat Tyr	cac His	aca Thr	ttt Phe	gga Gly 640	agg Arg	cat His	aat Asn	2335
ttt Phe	aaa Lys 645	aag Lys	acc Thr	aca Thr	gca Ala	aac Asn 650	ttg Leu	gat Asp	ttg Leu	ttc Phe	ctg Leu 655	agg Arg	aga Arg	ttt Phe	aat Asn	2383
gaa Glu 660	att Ile	cag Gln	ttt Phe	tgg Trp	gtc Val 665	gtc Val	act Thr	gag Glu	atc Ile	tgc Cys 670	ctt Leu	tgt Cys	tct Ser	cag Gln	ctc Leu 675	2431

agc aag cgt gtt cag cta tta aaa aaa ttt att aag ata gca gcc cac 2479
 Ser Lys Arg Val Gln Leu Leu Lys Lys Phe Ile Lys Ile Ala Ala His
 680 685 690

tgt aag gag tat aaa aat ctg aat tcc ttt ttt gcc atc gtc atg gga 2527
 Cys Lys Glu Tyr Lys Asn Leu Asn Ser Phe Phe Ala Ile Val Met Gly
 695 700 705

cta agt aac att gct gtg agc cgc ttg gca cta acg tgg gag aaa ctg 2575
 Leu Ser Asn Ile Ala Val Ser Arg Leu Ala Leu Thr Trp Glu Lys Leu
 710 715 720

cca agc aag ttc aag aag ttc tat gcg gag ttt gaa agt tta atg gac 2623
 Pro Ser Lys Phe Lys Lys Phe Tyr Ala Glu Phe Glu Ser Leu Met Asp
 725 730 735

cct tca agg aac cac agg gcc tac agg ctg aca gta gct aag ctg gaa 2671
 Pro Ser Arg Asn His Arg Ala Tyr Arg Leu Thr Val Ala Lys Leu Glu
 740 745 750 755

cct cct ctc atc ccc ttc atg cct ttg ctc att aaa gat atg aca ttt 2719
 Pro Pro Leu Ile Pro Phe Met Pro Leu Leu Ile Lys Asp Met Thr Phe
 760 765 770

act cat gag ggg aac aag acg ttc att gac aat cta gta aac ttt gaa 2767
 Thr His Glu Gly Asn Lys Thr Phe Ile Asp Asn Leu Val Asn Phe Glu
 775 780 785

aaa atg cgc atg att gca aat acg gcc aga aca gtg aga tac tac agg 2815
 Lys Met Arg Met Ile Ala Asn Thr Ala Arg Thr Val Arg Tyr Tyr Arg
 790 795 800

agc caa ccc ttc aat cct gat gca gct caa gct aat aag aac cat cag 2863
 Ser Gln Pro Phe Asn Pro Asp Ala Ala Gln Ala Asn Lys Asn His Gln
 805 810 815

gat gtc cgg agt tat gta cgg caa tta aat gtg att gac aac cag aga 2911
 Asp Val Arg Ser Tyr Val Arg Gln Leu Asn Val Ile Asp Asn Gln Arg
 820 825 830 835

act tta tca cag atg tca cac aga tta gag cct cgt cga cca 2953
 Thr Leu Ser Gln Met Ser His Arg Leu Glu Pro Arg Arg Pro
 840 845

tagacatttc aaatgcccaa agcaacagtt tgtctccagt ccacaatttt caaaaatgcc 3013

<210> 18
 <211> 849
 <212> PRT
 <213> Homo sapiens

<400> 18
 Met Val Leu Trp Lys Arg Ala Leu Thr Met Thr Gly Phe Leu Thr Arg
 1 5 10 15

Arg Thr His Leu Ile Glu Pro His Val Pro Leu Arg Pro Ala Asn Thr
 20 25 30

Ile Thr Lys Val Pro Ser Glu Lys Ile Leu Arg Ala Gly Lys Ile Leu
 35 40 45

Arg Asn Ala Ile Leu Ser Arg Ala Pro His Met Ile Arg Asp Arg Lys
 50 55 60

Tyr His Leu Lys Thr Tyr Arg Gln Cys Cys Val Gly Thr Glu Leu Val
 65 70 75 80

Asp Trp Met Ile Asp Glu Thr Pro Cys Val His Ser Arg Thr Gln Ala

85										90					95				
Val	Gly	Met	Trp	Gln	Val	Leu	Leu	Glu	Asp	Gly	Val	Leu	Asn	His	Val				
			100					105					110						
Asp	Gln	Glu	His	His	Phe	Gln	Asp	Phe	Tyr	Leu	Phe	Tyr	Arg	Phe	Leu				
		115					120					125							
Asp	Asp	Glu	His	Glu	Asp	Ala	Pro	Leu	Pro	Thr	Glu	Glu	Glu	Lys	Lys				
	130					135					140								
Glu	Cys	Asp	Glu	Glu	Leu	Gln	Asp	Thr	Met	Leu	Leu	Leu	Ser	Gln	Met				
145					150					155					160				
Gly	Pro	Asp	Ala	His	Met	Arg	Met	Ile	Leu	Arg	Lys	Pro	Pro	Gly	Gln				
				165					170					175					
Arg	Thr	Val	Asp	Asp	Leu	Glu	Ile	Ile	Tyr	Glu	Glu	Leu	Leu	His	Ile				
			180					185					190						
Lys	Ala	Leu	Ser	His	Leu	Ser	Thr	Thr	Val	Lys	Arg	Glu	Leu	Ala	Gly				
	195						200					205							
Val	Leu	Ile	Phe	Glu	Ser	His	Ala	Lys	Gly	Gly	Thr	Val	Leu	Phe	Asn				
	210					215					220								
Gln	Gly	Glu	Glu	Gly	Thr	Ser	Trp	Tyr	Ile	Ile	Leu	Lys	Gly	Ser	Val				
225					230					235					240				
Asn	Val	Val	Ile	Tyr	Gly	Lys	Gly	Val	Val	Cys	Thr	Leu	His	Glu	Gly				
				245					250					255					
Asp	Asp	Phe	Gly	Lys	Leu	Ala	Leu	Val	Asn	Asp	Ala	Pro	Arg	Ala	Ala				
			260					265					270						
Ser	Ile	Val	Leu	Arg	Glu	Asp	Asn	Cys	His	Phe	Leu	Arg	Val	Asp	Lys				
		275					280					285							
Glu	Asp	Phe	Asn	Arg	Ile	Leu	Arg	Asp	Val	Glu	Ala	Asn	Thr	Val	Arg				
	290					295					300								
Leu	Lys	Glu	His	Asp	Gln	Asp	Val	Leu	Val	Leu	Glu	Lys	Val	Pro	Ala				
305					310					315					320				
Gly	Asn	Arg	Ala	Ser	Asn	Gln	Gly	Asn	Ser	Gln	Pro	Gln	Gln	Lys	Tyr				
				325					330					335					
Thr	Val	Met	Ser	Gly	Thr	Pro	Glu	Lys	Ile	Leu	Glu	His	Phe	Leu	Glu				
			340					345					350						
Thr	Ile	Arg	Leu	Glu	Ala	Thr	Leu	Asn	Glu	Ala	Thr	Asp	Ser	Val	Leu				
		355					360					365							
Asn	Asp	Phe	Ile	Met	Met	His	Cys	Val	Phe	Met	Pro	Asn	Thr	Gln	Leu				
	370					375					380								
Cys	Pro	Ala	Leu	Val	Ala	His	Tyr	His	Ala	Gln	Pro	Ser	Gln	Gly	Thr				
385					390					395					400				
Glu	Gln	Glu	Lys	Met	Asp	Tyr	Ala	Leu	Asn	Asn	Lys	Arg	Arg	Val	Ile				
				405					410					415					
Arg	Leu	Val	Leu	Gln	Trp	Ala	Ala	Met	Tyr	Gly	Asp	Leu	Leu	Gln	Glu				
			420					425					430		</				

450					455					460					
Glu	Lys	Ile	Val	Lys	Gln	Ile	Ser	Glu	Asp	Ala	Lys	Ala	Pro	Gln	Lys
465					470					475					480
Lys	His	Lys	Val	Leu	Leu	Gln	Gln	Phe	Asn	Thr	Gly	Asp	Glu	Arg	Ala
				485					490					495	
Gln	Lys	Arg	Gln	Pro	Ile	Arg	Gly	Ser	Asp	Glu	Val	Leu	Phe	Lys	Val
			500					505					510		
Tyr	Cys	Met	Asp	His	Thr	Tyr	Thr	Thr	Ile	Arg	Val	Pro	Val	Ala	Thr
		515					520					525			
Ser	Val	Lys	Glu	Val	Ile	Ser	Ala	Val	Ala	Asp	Lys	Leu	Gly	Ser	Gly
		530				535					540				
Glu	Gly	Leu	Ile	Ile	Val	Lys	Met	Ser	Ser	Gly	Gly	Glu	Lys	Val	Val
545				550						555					560
Leu	Lys	Pro	Asn	Asp	Val	Ser	Val	Phe	Thr	Thr	Leu	Thr	Ile	Asn	Gly
				565					570					575	
Arg	Leu	Phe	Ala	Cys	Pro	Arg	Glu	Gln	Phe	Asp	Ser	Leu	Thr	Pro	Leu
			580					585					590		
Pro	Glu	Gln	Glu	Gly	Pro	Thr	Val	Gly	Thr	Val	Gly	Thr	Phe	Glu	Leu
		595					600					605			
Met	Ser	Ser	Lys	Asp	Leu	Ala	Tyr	Gln	Met	Thr	Ile	Tyr	Asp	Trp	Glu
	610					615					620				
Leu	Phe	Asn	Cys	Val	His	Glu	Leu	Glu	Leu	Ile	Tyr	His	Thr	Phe	Gly
625						630					635				640
Arg	His	Asn	Phe	Lys	Lys	Thr	Thr	Ala	Asn	Leu	Asp	Leu	Phe	Leu	Arg
				645					650					655	
Arg	Phe	Asn	Glu	Ile	Gln	Phe	Trp	Val	Val	Thr	Glu	Ile	Cys	Leu	Cys
			660					665					670		
Ser	Gln	Leu	Ser	Lys	Arg	Val	Gln	Leu	Leu	Lys	Lys	Phe	Ile	Lys	Ile
		675				680						685			
Ala	Ala	His	Cys	Lys	Glu	Tyr	Lys	Asn	Leu	Asn	Ser	Phe	Phe	Ala	Ile
		690				695					700				
Val	Met	Gly	Leu	Ser	Asn	Ile	Ala	Val	Ser	Arg	Leu	Ala	Leu	Thr	Trp
705						710					715				720
Glu	Lys	Leu	Pro	Ser	Lys	Phe	Lys	Lys	Phe	Tyr	Ala	Glu	Phe	Glu	Ser
				725					730					735	
Leu	Met	Asp	Pro	Ser	Arg	Asn	His	Arg	Ala	Tyr	Arg	Leu	Thr	Val	Ala
			740					745					750		
Lys	Leu	Glu	Pro	Pro	Leu	Ile	Pro	Phe	Met	Pro	Leu	Leu	Ile	Lys	Asp
		755				760					765				
Met	Thr	Phe	Thr	His	Glu	Gly	Asn	Lys	Thr	Phe	Ile	Asp	Asn	Leu	Val
	770					775					780				
Asn	Phe	Glu	Lys	Met	Arg	Met	Ile	Ala	Asn	Thr	Ala	Arg	Thr	Val	Arg
				790							795				800
Tyr	Tyr	Arg	Ser	Gln	Pro	Phe	Asn	Pro	Asp	Ala	Ala	Gln	Ala	Asn	Lys
				805					810					815	
Asn	His	Gln	Asp	Val	Arg	Ser	Tyr	Val	Arg	Gln	Leu	Asn	Val	Ile	Asp
